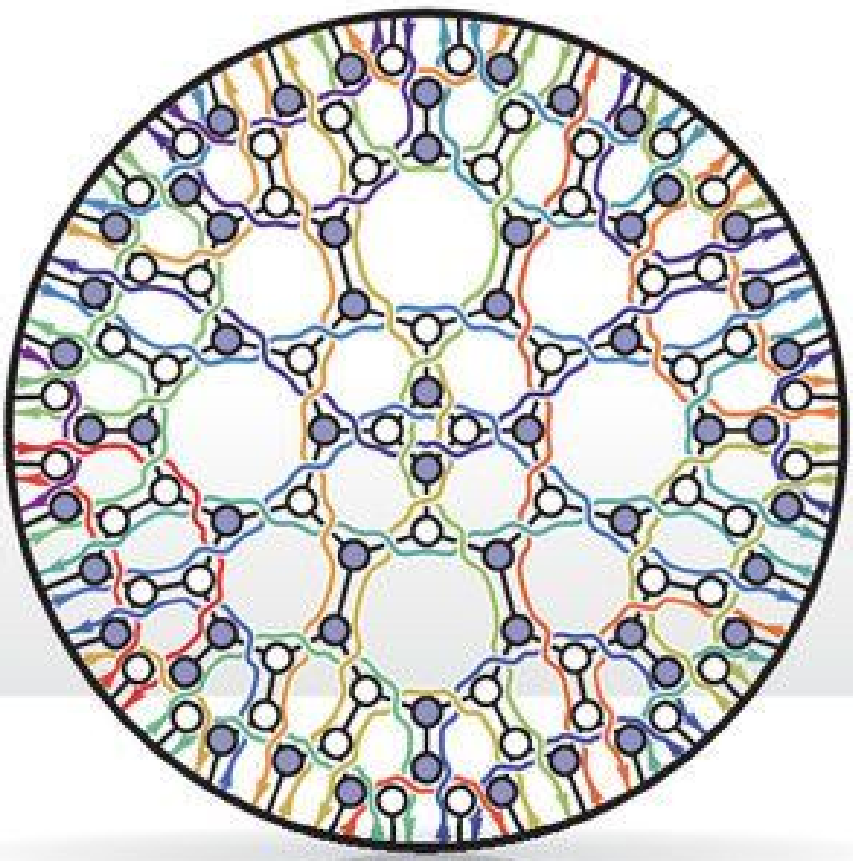
NIMA **ARKANI-HAMED**
JACOB **BOURJAILY**
FREDDY **CACHAZO**
ALEXANDER **GONCHAROV**
ALEXANDER **POSTNIKOV**
JAROSLAV **TRNKA**

GRASSMANNIAN GEOMETRY OF SCATTERING AMPLITUDES



GRASSMANNIAN GEOMETRY OF SCATTERING AMPLITUDES

Outlining a revolutionary reformulation of the foundations of perturbative quantum field theory, this book is a self-contained and authoritative analysis of the application of this new formulation to the case of planar, maximally supersymmetric Yang-Mills theory. The book begins by deriving connections between scattering amplitudes and Grassmannian geometry from first principles before introducing novel physical and mathematical ideas in a systematic manner accessible to both physicists and mathematicians. The principle players in this process are on-shell functions which are closely related to certain sub-strata of Grassmannian manifolds called positroids—in terms of which the classification of on-shell functions and their relations becomes combinatorially manifest. This is an essential introduction to the geometry and combinatorics of the positroid stratification of the Grassmannian and an ideal text for advanced students and researchers working in the areas of field theory, high energy physics, and the broader fields of mathematical physics.

Nima Arkani-Hamed is Professor of Physics at the Institute for Advanced Study, Princeton, New Jersey, USA.

Jacob Bourjaily is Assistant Professor of Physics at the Niels Bohr International Academy and Discovery Center at the University of Copenhagen, Copenhagen, Denmark.

Freddy Cachazo is the Gluskin Sheff Freeman Dyson Chair in Theoretical Physics at the Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada.

Alexander Goncharov is Professor of Mathematics at Yale University, New Haven, Connecticut, USA.

Alexander Postnikov is Professor of Applied Mathematics and Algebraic Combinatorics at the Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Jaroslav Trnka is Assistant Professor of Physics at the University of California, Davis, California, USA.

GRASSMANNIAN GEOMETRY OF SCATTERING AMPLITUDES

NIMA ARKANI-HAMED
Institute for Advanced Study

JACOB BOURJAILY
Niels Bohr Institute

FREDDY CACHAZO
Perimeter Institute

ALEXANDER GONCHAROV
Yale University

ALEXANDER POSTNIKOV
Massachusetts Institute of Technology

JAROSLAV TRNKA
University of California, Davis

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107086586

First published 2016

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Names: Arkani-Hamed, Nima.
Title: Grassmannian geometry of scattering amplitudes / Nima Arkani-Hamed (Institute for Advanced Study) [and five others].
Description: New York : Cambridge University Press, 2016. | Includes bibliographical references and index.
Identifiers: LCCN 2015046557 | ISBN 9781107086586 (hardback)
Subjects: LCSH: Quantum field theory. | Particles (Nuclear physics) | Graph theory. | Combinatorial analysis. | Geometry, Algebraic.
Classification: LCC QC174.45.G67 2016 | DDC 516.3/5–dc23
LC record available at http://lccn.loc.gov/2015046557

ISBN 978-1-107-08658-6 Hardback

Contents

Acknowledgments *page* ix

1 Introduction 1

2 Introduction to on-shell functions and diagrams 5
- 2.1 On-shell particles, functions, and kinematical data 5
- 2.2 Scattering amplitudes and their amalgamations 8
- 2.3 On-shell building blocks: the massless three-particle S-matrix 10
- 2.4 On-shell supersymmetry and (maximally) supersymmetric theories 14
- 2.5 Building up diagrams with "BCFW-bridges" 18
- 2.6 On-shell recursion for all-loop amplitudes 20
- 2.7 Physical equivalences among on-shell diagrams 23

3 Permutations and scattering amplitudes 29
- 3.1 Combinatorial descriptions of scattering processes 29
- 3.2 The BCFW-bridge construction of representative graphs 33

4 From on-shell diagrams to the Grassmannian 37
- 4.1 The Grassmannian of k-planes in n dimensions, $G(k,n)$ 37
- 4.2 Grassmannian description of kinematical data—λ and $\tilde{\lambda}$ 40
- 4.3 Grassmannian representation of on-shell diagrams 41
- 4.4 Amalgamation of on-shell diagrams 45
- 4.5 "Boundary measurements" and canonical coordinates 49
- 4.6 Coordinate transformations induced by moves and reduction 52
- 4.7 Relation to composite leading singularities 56

5 Configurations of vectors and the *positive* Grassmannian 60
- 5.1 The geometry and combinatorics of the positroid stratification 60
- 5.2 Canonical coordinates and the equivalence of permutation labeling 66

5.3 Positroid cells and the positive part of the Grassmannian 69
5.4 Canonically positive coordinates for positroids 73

6 Boundary configurations, graphs, and permutations 77
6.1 Physical singularities and positroid boundaries 77
6.2 Boundary configurations: combinatorics and stratification 78
6.3 (Combinatorial) polytopes in the Grassmannian 79
6.4 Approaching boundaries in canonical coordinates 81

7 The invariant top-form and the positroid stratification 83
7.1 Equivalence with the canonical positroid volume form 85

8 (Super-)conformal and dual conformal invariance 88
8.1 The Grassmannian geometry of momentum conservation 88
8.2 Twistor space and the superconformality of on-shell forms 90
8.3 Momentum-twistors and dual super-conformal invariance 92

9 Positive diffeomorphisms and Yangian invariance 97

10 The kinematical support of physical on-shell forms 101
10.1 Kinematical support of NMHV Yangian-invariants 102
10.2 Kinematical support for one-dimensional kinematics 103
10.3 A general combinatorial test of kinematical support 103

11 Homological identities among Yangian-invariants 107
11.1 Homological identities in the Grassmannian 108

12 (Relatively) orienting canonical coordinate charts on positroids configurations 113
12.1 Comparing the orientations of canonical coordinate charts 115
12.2 Accessing boundary configurations from the canonical atlas 118

13 Classification of Yangian-invariants and their relations 121

14 The Yang–Baxter relation and ABJM theories 127
14.1 The on-shell avatar of the Yang–Baxter relation 127
14.2 ABJM theories 130

15 On-shell diagrams for theories with $\mathcal{N} < 4$ supersymmetries 136

16 Dual graphs and cluster algebras 141
16.1 The 'dual' of an on-shell diagram 141
16.2 Cluster algebras: seeds, mutations, and cluster coordinates 146
16.3 Cluster amalgamation 150
16.4 (Brief) overview of cluster structures in physics 152

17 On-shell representations of scattering amplitudes 155
17.1 (Diagrammatic) proof of the BCFW recursion relations 156
17.2 The structure of (tree) amplitudes in the Grassmannian 159

17.3 Canonical coordinates for loop integrands 164
17.4 The transcendentality of loop amplitudes 173

18 Outlook 178

References 183

Index 192

Acknowledgments

This book represents the synthesis of several years of work and was originally made available on the arXiv as reference [1]. There are many people without whom this book would not have been possible, or without whose comments would not be nearly as comprehensive as we hope that it is. Many of the mathematical topics discussed were greatly improved through early discussions with Pierre Deligne, Bob MacPherson, and Mark Goresky at the Institute for Advanced Study; and for the physics topics discussed, we owe an enormous debt of gratitude for discussions with Andrew Hodges and Simon Caron-Huot.

For helpful comments regarding early versions of the manuscript, we are grateful to Simon Caron-Huot, Henriette Elvang, Sebastian Franco, Yu-tin Huang, Lionel Mason, Marcus Spradlin, Cristian Vergu, Congkao Wen, and Lauren Williams.

This work was made possible through the hospitality of the School of Natural Sciences at the Institute for Advanced Study, and was supported in part by grants from the National Science Foundation DMS-1100147 (A.P.), DMS-1059129 (A.G.), and PHY-0756966 (J.T.), the Harvard Society of Fellows and a grant from the Harvard Milton Fund (J.B.), the NSERC of Canada and MEDT of Ontario (F.C.), and U.S. Department of Energy contracts DE-FG02-91ER40654 (N.A.-H.) and DE-FG02-91ER40654 (J.B.). The authors are also grateful to the Aspen Public Library for their hospitality.

1

Introduction

The traditional formulation of quantum field theory—encoded in its very name—is built on the two pillars of *locality* and *unitarity* [2]. The standard apparatus of Lagrangians and path integrals allows us to make these two fundamental principles manifest. This approach, however, requires the introduction of a large amount of unphysical redundancy in our description of physical processes. Even for the simplest case of scalar field theories, there is the freedom to perform field redefinitions. Starting with massless particles of spin 1 or higher, we are forced to introduce even larger, gauge redundancies [2].

Over the past few decades, there has been a growing realization that these redundancies hide amazing physical and mathematical structures lurking within the heart of quantum field theory. This has been seen dramatically at strong coupling in gauge/gauge (see, e.g. [3–5]) and gauge/gravity dualities [6]. The past decade has uncovered further remarkable new structures in field theory even at weak coupling, seen in the properties of scattering amplitudes in gauge theories and gravity (for reviews, see [7–12]). The study of scattering amplitudes is fundamental to our understanding of field theory, and fueled its early development in the hands of Feynman, Dyson, and Schwinger among others. It is therefore surprising to see that even here, by committing so strongly to particular, gauge-redundant descriptions of the physics, the usual formalism is completely blind to astonishingly simple and beautiful properties of the gauge-invariant physical observables of the theory.

Many of the recent developments have been driven by an intensive exploration of $\mathcal{N}=4$ supersymmetric Yang–Mills (SYM) in the planar limit [12, 13]. The all-loop integrand for scattering amplitudes in this theory can be determined by a generalization of the BCFW recursion relations [14], in a way that is closely tied to remarkable new structures in algebraic geometry, associated with contour integrals over the Grassmannian $G(k,n)$ [15–18]. This makes both the *conformal* and long-hidden *dual conformal* invariance of the theory (which together close into the infinite-dimensional Yangian symmetry) completely manifest [19]. It

is remarkable that a single function of external kinematical variables can be interpreted as a scattering amplitude in one space-time, and as a Wilson loop in another (for a review, see [12]). Each of these descriptions makes a commitment to locality in its own space-time, making it impossible to see the dual picture. By contrast, the Grassmannian picture makes no mention of locality or unitarity, and does not commit to any gauge-redundant description of the physics, allowing it to manifest *all* the symmetries of the theory.

There has also been extraordinary progress in determining the amplitude itself beyond the integrand, using the technology of symbols of transcendental functions to powerfully constrain and control the polylogarithms occurring in the final results [20, 21]. While a global picture is still missing, a huge amount of data has been generated. The symbol for all 2-loop MHV amplitudes has been determined [22] (see also [23]), and a handful of 2-loop NMHV and 3-loop MHV symbols have been found [24–26]. Remarkable strategies have also been presented to bootstrap amplitudes to very high loop orders [27–31]. Many of these ideas have a strong resonance with the explosion of progress in the last decade using integrability to find exact results in planar $\mathcal{N} = 4$, starting with the spectacular solution of the spectral problem for anomalous dimensions [13, 32].

All of these developments have made it completely clear that there are powerful new mathematical structures underlying the extraordinary properties of scattering amplitudes in gauge theories. If history is any guide, formulating and understanding the physics in a way that makes the symmetries manifest should play a central role in the story. The Grassmannian picture does this, but up to this point there has been little understanding as to why this formulation exists, exactly how it works, or where it comes from physically. Our primary goal in this note is to resolve this unsatisfactory state of affairs.

We will derive the connection between scattering amplitudes and the Grassmannian, starting physically from first principles. This will lead us into direct contact with several beautiful and active areas of current research in mathematics [33–42]. The past few decades have seen vigorous interactions between physics and mathematics in a wide variety of areas, but what is going on here involves *new* areas of mathematics that have only very recently played any role in physics, involving simple but deep ideas ranging from combinatorics to algebraic geometry. It is both startling and exciting that such elementary mathematical notions are found at the heart of the physics of scattering amplitudes.

This new way of thinking about scattering amplitudes involves many novel physical and mathematical ideas. Our presentation will be systematic, and we have endeavored to make it self contained and completely accessible to physicists. While we will discuss a number of mathematical results—some of them new—we will usually be content with the physicist's level of rigor. While the essential ideas here are all very simple, they are tightly interlocking, and range over a wide variety

of areas—most of which are unfamiliar to most physicists. Thus, before jumping into the detailed exposition, as a guide to the reader we end this introductory chapter by giving a roadmap of the logical structure and content of the book.

In Chapter 2, we introduce the central physical idea motivating our work, which is to focus on *on-shell diagrams*, obtained by gluing together fundamental 3-particle amplitudes and integrating over the on-shell phase space of internal particles. These objects are of central importance to the understanding scattering amplitudes. We will see that scattering amplitudes in planar $\mathcal{N}=4$—to all loop orders—can be represented *directly* in terms of on-shell processes. In this picture, "virtual particles" make no appearance at all. We should emphasize that we are not merely using on-shell information to determine scattering amplitudes, but rather seeing that the amplitudes can be directly computed in terms of fully on-shell processes. The off-shell, virtual particles familiar from Feynman diagrams are replaced by internal, *on-shell* particles (with generally complex momenta).

In our study of on-shell diagrams, we will see that different diagrams related by certain elementary moves can be physically equivalent, leading to the natural question of how to invariantly characterize their physical content. Remarkably, the invariant content of on-shell diagrams turns out to be characterized by *combinatorial* data. We discuss this in detail in Chapter 3 where we show how a long-known and beautiful connection between permutations and scattering amplitudes in integrable $(1+1)$-dimensional theories generalizes to more realistic theories in $(3+1)$ dimensions.

In Chapter 4 we turn to actually calculating on-shell diagrams and find that the most natural way of carrying out the computations is to associate each diagram with a certain differential form on an *auxiliary* Grassmannian. In Chapter 5 and 6 we show how the invariant, combinatorial content of an on-shell diagram is reflected in the Grassmannian directly. This is described in terms of a surprisingly simple stratification of the configurations of k-dimensional vectors endowed with a cyclic ordering, classified by the linear dependencies among consecutive chains of vectors. For the real Grassmannian, this stratification can be equivalently described in an amazingly simple and beautiful way as nested 'boundaries' of the *positive part* of the Grassmannian [33], which is motivated by the theory of totally positive matrices [34, 43, 44]. Each on-shell diagram can then be associated with a particular configuration or "stratum" among the boundaries of the positive Grassmannian.

In Chapter 7 we make contact with the Grassmannian contour integral of reference [15], which is now seen as a compact way of representing the natural, invariant top-form on the positive Grassmannian. This form of the measure allows us to easily identify the conformal and dual conformal symmetries of the theory, which are related by a simple mapping of permutations described in Chapter 8. In Chapter 9, we show that the invariance of scattering amplitudes under the action

of the level-one generators of the Yangian has a transparent interpretation: these generators correspond to the leading nontrivial diffeomorphisms that preserve all the cells of the positive Grassmannian.

In Chapter 10 we begin a systematic classification of Yangian-invariants and their relations by first describing a combinatorial test to determine whether an on-shell diagram has non-vanishing kinematical support (and if it has support, how many solutions exist). In Chapter 11 a geometric basis is given for all the myriad, highly nontrivial identities satisfied among Yangian-invariants. This completes the classification of *all* Yangian-Invariants together with *all* their relations. In Chapter 13, we give a tour of this classification as it emerges through N^4MHV.

In Chapter 14 we show that the story for scattering amplitudes in integrable $(1+1)$-dimensional theories—in particular, the Yang–Baxter relation—can be understood as a special case of our general results regarding on-shell diagrams. We further show that scattering amplitudes for the ABJM theory in $(2+1)$ dimensions [45] can also be computed in terms of a natural specialization of on-shell diagrams: those associated with the null orthogonal Grassmannian. And we initiate the study of on-shell diagrams in theories with less (or no) supersymmetry in Chapter 15.

The positive Grassmannian is naturally endowed with a rich mathematical structure known as a *cluster algebra*—the original theory of which was developed in [35] and has since been generalized to the theory of *cluster varieties* in [37, 38]. Remarkably, this structure has made striking appearances in widely disparate parts of physics in the last decade—from conformal blocks for higher Toda theories [36, 46], to wall-crossing phenomena [47, 48], to quiver gauge theories with $\mathcal{N}=1$ super-conformal symmetry [49–54], to soliton solutions to the KP equation [55–57]. We briefly review this story in Chapter 16, and also summarize its various physical manifestations in hopes of stimulating a deeper understanding for these extremely surprising connections between physics and mathematics.

In Chapter 17 we move beyond the discussion of individual on-shell diagrams and describe the particular combinations that represent scattering amplitudes. We present a self-contained direct proof—using on-shell diagrams alone—that the BCFW construction of the all-loop integrand generates an object with precisely those singularities dictated by quantum field theory. We then show that the Grassmannian representation of loop integrands are always given in a remarkable "*d*log" form, which we illustrate using examples of simple one- and two-loop amplitudes. We discuss the implications of this representation for the transcendental functions that arise after the loop integrands are integrated.

We conclude our story in Chapter 18 with a discussion of a number of the outstanding open directions for further research.

2

Introduction to on-shell functions and diagrams

Theoretical explorations in field theory have been greatly advanced by focusing on interesting classes of observables—from local correlation functions and scattering amplitudes, to Wilson and 't Hooft loops, surface operators and line defects, to partition functions on various manifolds (see, e.g., [58, 59]). The central physical idea of our work is to extend the notion of "scattering amplitudes" to a broader class of objects called *on-shell functions*, which we introduce in this chapter.

2.1 On-shell particles, functions, and kinematical data

On-shell functions, like the S-matrix, depend only on the data describing physically observable external states. This data consists of the momentum $p_a^\mu \in \mathbb{R}^{3,1}$, mass m_a, spin σ_a, helicity $h_a \in \{\sigma_a, \sigma_a - 1, \ldots, -\sigma_a\}$, and any non-kinematical quantum numbers q_a that describe the external particle indexed by $a \in \{1, \ldots, n\}$. The momentum of any *observable* state satisfies the Einstein relation, $p^\mu p_\mu = m^2$; such particles are said to be "on the mass-shell" (the hyperboloid $p^\mu p_\mu - m^2 = 0$), or simply *on-shell*. In this work, we will focus on theories involving *massless* particles—those with $m_a = 0$; such particles can only have helicity $h_a = \pm\sigma_a$.

When an external particle is massless, the (2×2)-matrix constructed out of the Pauli matrices $\sigma_\mu^{\alpha\dot\alpha}$,

$$p_a^{\alpha\dot\alpha} \equiv p_a^\mu \sigma_\mu^{\alpha\dot\alpha} = \begin{pmatrix} p_a^0 + p_a^3 & p_a^1 - ip_a^2 \\ p_a^1 + ip_a^2 & p_a^0 - p_a^3 \end{pmatrix}, \tag{2.1}$$

(with entries labeled by $\alpha = 1, 2$, and $\dot\alpha = \dot 1, \dot 2$) will have a vanishing determinant:

$$\det(p_a^{\alpha\dot\alpha}) = (p_a^0)^2 - (p_a^1)^2 - (p_a^2)^2 - (p_a^3)^2 = 0 \ (\equiv \eta_{\mu\nu} p_a^\mu p_a^\nu); \tag{2.2}$$

and so $p_a^{\alpha\dot\alpha}$ has rank 1 (or 0). We can make this manifest via the substitution

$$p_a^{\alpha\dot\alpha} \equiv \lambda_a^\alpha \widetilde{\lambda}_a^{\dot\alpha} \Leftrightarrow \text{“}a\rangle[a\text{”}, \tag{2.3}$$

where $\lambda_a^\alpha, \widetilde{\lambda}_a^{\dot{\alpha}} \in \mathbb{C}^2$ are called *spinor-helicity* variables [60–64]. If the momentum p_a were *real*, we would have $\widetilde{\lambda}_a = \pm\lambda_a^*$; but we will often find it useful to allow all momenta to be complex, and consider $\lambda_a, \widetilde{\lambda}_a$ to be independent variables.

The rescaling $\lambda_a \mapsto t_a\lambda_a, \widetilde{\lambda}_a \mapsto t_a^{-1}\widetilde{\lambda}_a$ leaves the momentum p_a, (2.3), invariant and represents the action of the *little group* (for more details, see e.g. [2, 65]). Upon its complexification, the local Lorentz group becomes $SL(2)_L \times SL(2)_R$, with λ^α and $\widetilde{\lambda}^{\dot{\alpha}}$ transforming in the fundamental representations of $SL(2)_L$ and $SL(2)_R$, respectively. Knowing this, we may construct the Lorentz-invariants,

$$\langle a\,b\rangle \equiv \det\{\lambda_a, \lambda_b\} \quad \text{and} \quad [a\,b] \equiv \det\{\widetilde{\lambda}_a, \widetilde{\lambda}_b\}, \tag{2.4}$$

out of which all Lorentz-invariants of the momenta can be constructed—e.g.,

$$\eta_{\mu\nu}\,(p_a + p_b)^\mu\,(p_a + p_b)^\nu \equiv (p_a + p_b)^2 = \langle a\,b\rangle[a\,b], \tag{2.5}$$

for any on-shell, massless four-momenta p_a, p_b.

The Lorentz-invariant phase space ('LIPS') associated with an on-shell particle is given by the four degrees of freedom of $\lambda_a, \widetilde{\lambda}_a$ modulo the action of the little group—a $GL(1)$ redundancy. Thus, the differential form describing an on-shell particle's phase space can be written

$$d^3\text{LIPS}_a \equiv \frac{d^2\lambda_a\, d^2\widetilde{\lambda}_a}{\text{vol}(GL(1))}, \tag{2.6}$$

where "$1/\text{vol}(GL(1))$" represents the instruction to eliminate the $GL(1)$ redundancy of the little group, resulting in a three-dimensional form on phase space.

In general, we view all on-shell functions as being decorated with the Lorentz-invariant phase space measures for each external particle. As such, we may view them more formally as on-shell (differential) forms on the phase space of all the external kinematical data. Because of this, we will refer to on-shell functions interchangeably as on-shell forms throughout this work.

Let us denote the external wave function for particle a with helicity $h_a = \pm\sigma_a$ by '$|a\rangle^{h_a}$' (which should not be confused with the notation '$a\rangle$' for λ_a). Under the action of the little group, $\lambda_a \mapsto t_a\lambda_a, \widetilde{\lambda}_a \mapsto t_a^{-1}\widetilde{\lambda}_a$, the wave function transforms according to (see e.g. reference [2]):

$$|a\rangle^{h_a} \mapsto t_a^{-2h_a}|a\rangle^{h_a}. \tag{2.7}$$

Because of this, any Lorentz-invariant on-shell function of the external states (e.g. a scattering amplitude) must transform under the little group accordingly:

$$f(\ldots, t_a\lambda_a, t_a^{-1}\widetilde{\lambda}_a, h_a, \ldots) = t_a^{-2h_a} f(\ldots, \lambda_a, \widetilde{\lambda}_a, h_a, \ldots). \tag{2.8}$$

We will see in section 2.3 that this scaling property together with momentum conservation uniquely fixes the kinematical dependence of the S-matrix for three massless particles with any helicities (to all loop orders!).

Before moving on to the main subject of this chapter, on-shell diagrams, it will be helpful to establish some useful notational conventions. As we have seen, the kinematical data describing on-shell particles are specified by a pair of two-vectors $\lambda_a, \widetilde{\lambda}_a \in \mathbb{C}^2$ for each particle $a \in \{1,\ldots,n\}$. We will frequently find it convenient to organize these data collectively into a pair of $(2\times n)$ matrices written according to the following convention:

$$\{\lambda_a^\alpha\} \Leftrightarrow \lambda \equiv \begin{pmatrix} \lambda_1^1 \cdots \lambda_n^1 \\ \lambda_1^2 \cdots \lambda_n^2 \end{pmatrix} \equiv \begin{pmatrix} \lambda^1 \\ \lambda^2 \end{pmatrix} \equiv \big(\lambda_1 \cdots \lambda_n\big), \tag{2.9}$$

and similarly for $\widetilde{\lambda}$. That is, 'λ' collectively denotes all external spinors and components $\{\lambda_a^\alpha\}$, 'λ_a' denotes the two-vector of spinor components for the ath particle, and 'λ^α' denotes the n-vector consisting of each spinor's α-component.

Finally, momentum conservation must always be satisfied for the particles involved in any scattering process. If we conventionally take all the external momenta to be incoming, then momentum conservation becomes the constraint

$$\delta^{2\times 2}\left(\sum_{a=1}^{n} \lambda_a^\alpha \widetilde{\lambda}_a^{\dot\alpha}\right) \equiv \delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big), \tag{2.10}$$

where we have introduced "$\cdot$" to denote a summation over external particles. This will prove a useful convention throughout the rest of this work. Notice also that because the four-momenta are each written as a (2×2) matrix according to (2.1), momentum conservation is naturally organized into a (2×2) matrix of constraints. Whenever a system of constraints is naturally organized into a $(k\times m)$ matrix, we will indicate this by writing "$\delta^{k\times m}\big(\cdots\big)$" as in (2.10).

We should clarify that, throughout this work, we consider only "holomorphic" δ-functions. These behave very similarly to the traditional "δ-functions" of quantum mechanics, but require no use of a complex norm. Concretely, our δ-functions are defined as *residue* (or *contour*) prescriptions according to:

$$\int dz\, g(z)\, \delta\big(f(z)\big) \equiv \sum_{z^*|f(z^*)=0} \mathrm{Res}\left(\frac{g(z)}{f(z)}; z^*\right). \tag{2.11}$$

The following is a simple, illustrative example:

$$\int dz\, \delta\big(f(z)\big) = \sum_{z^*|f(z^*)=0} \frac{1}{f'(z^*)}, \tag{2.12}$$

where $f'(z^*)$ denotes the derivative of f with respect to z, evaluated at z^*. The reader should notice that δ-functions defined in this way behave the same way as ordinary δ-functions, except that, in (2.12) for example, no absolute-value sign appears around the derivative of f. This definition generalizes to the case $\delta^{k\times m}\big(\cdots\big)$, which specifies the contour for a co-dimension $(k\times m)$ residue (see e.g. [66]).

Thus, if we re-write the four components of the δ-functions encoding momentum conservation, (2.10), in terms of Lorentz-invariant quantities by contracting

with some reference null momenta $\{\lambda_r\widetilde{\lambda}_r, \lambda_s\widetilde{\lambda}_s\}$, we pick-up an ordinary *Jacobian* from the change of variables (without any absolute-value signs!):

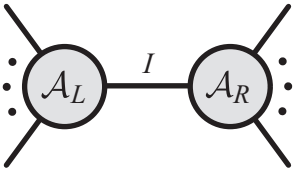
$$\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big) = \langle rs\rangle[rs]\,\delta\Big(\sum_a \langle ra\rangle[ar]\Big)\delta\Big(\sum_a \langle ra\rangle[as]\Big)\delta\Big(\sum_a \langle sa\rangle[ar]\Big)\delta\Big(\sum_a \langle sa\rangle[as]\Big). \tag{2.13}$$

2.2 Scattering amplitudes and their amalgamations

We would like to understand the entire class of physically meaningful functions describable *exclusively* in terms of on-shell kinematical data for some number of external states (without any reference to virtual particles, gauge redundancies, ghosts, or any of the other unphysical baggage associated with the traditional approach to quantum field theory).

A particularly important example of this class of functions is the full scattering amplitude (the "*S*-matrix") $\mathcal{A}_n(\lambda,\widetilde{\lambda},h)$ for n particles with momenta given by $\lambda,\widetilde{\lambda}$ and with helicities $h \equiv \big(h_1\cdots h_n\big)$. But scattering amplitudes represent only a small subset of the meaningful gauge-invariant functions that can be constructed exclusively in terms of on-shell external data. In particular, knowing even a few scattering amplitudes, we can "amalgamate" them into more complex objects we call *on-shell functions*, which can be represented as *on-shell diagrams*, constructed out of amplitudes by sewing them together in a natural way.

The most familiar example of an on-shell function built out of scattering amplitudes, but which is not a scattering amplitude itself, is known as a *factorization channel*; diagrammatically, we represent a factorization channel as follows



(2.14)

A factorization channel is well defined without any reference to any off-shell degrees of freedom, and corresponds to the particular on-shell function

$$\sum_{h_I}\int \frac{d^2\lambda_I\, d^2\widetilde{\lambda}_I}{\mathrm{vol}(GL(1))}\mathcal{A}_L(\ldots,\lambda_I,\widetilde{\lambda}_I,h_I)\mathcal{A}_R(\lambda_I,-\widetilde{\lambda}_I,-h_I,\ldots). \tag{2.15}$$

(Here, we have left implicit a summation over any non-kinematical quantum numbers labeling the internal particle.) Notice that in (2.15), integration over the internal particle's on-shell phase space is trivial: it is *completely localized* by the momentum-conserving δ-functions present in the two amplitudes. Notice, however, that even after localizing the three-dimensional phase space integral, five of the eight initial δ-functions remain; these impose the four constraints of overall momentum conservation, together with the additional constraint $\big(\sum_{a\in L} p_a\big)^2 = 0$.

We can understand how the formula (2.15) follows from locality and unitarity as follows. Because the internal particle is massless (and on-shell), the two scattering amplitudes can involve particles separated by arbitrary distances in space and time; as such, locality dictates that they must be independent, and so the amplitudes must be multiplied together. And unitarity instructs us to marginalize over any unobserved states—integrating over each internal particle's on-shell phase space, (2.6), and summing over possible helicities (and any other quantum numbers).

This rule can be generalized to define arbitrary graphs built out of on-shell scattering amplitudes separated by internal (but on-shell) particles. Thus, the on-shell function associated with a graph Γ involving an amplitude $\mathcal{A}_v$ at each vertex v and any number of internal particles $i \in I$ is defined according to

$$f_\Gamma \equiv \prod_{i\in I}\left(\sum_{h_i}\int \frac{d^2\lambda_i d^2\widetilde{\lambda}_i}{\mathrm{vol}(GL(1))}\right)\prod_v \mathcal{A}_v, \tag{2.16}$$

where convention that amplitudes involve only *incoming* momenta dictates that two ends of any internal line must involve opposite momenta and helicities.

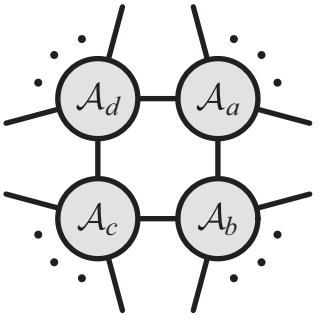
One important physical characteristic of any on-shell diagram obtained in the way described above is the number n_δ of constraints (if any) that are imposed on the external kinematical data beyond overall momentum conservation. Because each vertex amplitude imposes momentum conservation, we start with a total $4n_V$ constraints for a diagram with n_V vertices; these constraints *always* imply overall momentum conservation, and so $4n_V - 4$ of the constraints can contribute to n_δ. But we must integrate over each of the n_I internal particles' three dimensional on-shell phase space; therefore, if as many of these phase space integrals can be localized as possible, then the number of excess constraints would be given by

$$n_\delta \equiv 4n_V - 3n_I - 4. \tag{2.17}$$

When $n_\delta = 0$, it is possible that all the phase space integrals can be localized by the δ-functions, resulting in an on-shell diagram that neither imposes any constraints on the external kinematics, nor leaves us with any remaining integrals over phase space to perform. When this is the case, the on-shell diagram is simply a (rational) function of the external kinematical variables; such on-shell functions have historically been called "leading singularities" in the physics literature [12,67], for reasons we will discuss momentarily. When $n_\delta > 0$, however, the resulting on-shell function imposes some number of excess constraints on the external kinematics; these objects have historically been called "singularities" or said to have "singular support". Finally, when $n_\delta < 0$, there are internal phase space integrations that *cannot* be localized by the δ-functions, requiring that we specify a choice of contour for their integration. We will see in section 2.6 that with the appropriate choice of contours, these on-shell phase space integrations *exactly* reproduce the "loop-level" contributions in the Feynman expansion (where, in a

Feynman diagram, this integration would be performed over degrees of freedom associated with the *off-shell* "loop momenta" of *virtual* particles).

We have already seen an example of a diagram for which $n_\delta > 0$: the factorization channel, (2.14). The following is an example of a diagram with $n_\delta = 0$:



(2.18)

For this diagram, *all* the internal phase space integrations are localized by momentum conservation at the vertices; and so, (2.18) involves no internal degrees of freedom and imposes no additional constraints on the external kinematical data. Because of this, the diagram (2.18) represents an ordinary (rational) function of the external momenta (and has nothing directly to do with a "loop" in the sense of traditional, off-shell perturbation theory).

As mentioned earlier, the diagram (2.18) would have historically been called a "one-loop leading singularity"; this is because it can be interpreted as a co-dimension–four residue of the one-loop Feynman integrand—where each residue constrains an *off-shell* internal (hence virtual) particle to be put *on-shell*. We choose not to use this terminology here, because it subordinates on-shell functions like (2.18) relative to objects such as Feynman diagrams defined in terms of unphysical, off-shell degrees of freedom. As emphasized above, *all* on-shell diagrams are intrinsically well defined, without any reference to Feynman diagrams.

2.3 On-shell building blocks: the massless three-particle *S*-matrix

As we saw in section 2.2, on-shell scattering amplitudes can be "amalgamated" into more complicated on-shell objects; but we cannot begin to investigate such objects until we know some scattering amplitudes to feed into the machinery. In this section, we show that the *S*-matrix involving three massless particles (with arbitrary helicities) is uniquely fixed by first principles *to all orders of perturbation theory*. Given just these basic amplitudes as building blocks, the procedure described above immediately leads to an incredible variety of on-shell diagrams and corresponding functions, defined to all orders of perturbation theory. Moreover, as we will see in section 2.6, the entire perturbative *S*-matrix (at least in

planar, maximally supersymmetric ('$\mathcal{N}=4$') Yang–Mills theory) can be expressed directly in terms of these diagrams alone.

Three-particle amplitudes are the fundamental building blocks from which everything else can be constructed. Let us now see how they are in fact completely determined (up to an overall coupling constant) by Poincaré invariance—momentum conservation and little group scaling. This is a consequence of the unique simplicity of three-particle kinematics.

For three massless particles, momentum conservation is the constraint,

$$\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big) \implies \lambda_1\widetilde{\lambda}_1+\lambda_2\widetilde{\lambda}_2+\lambda_3\widetilde{\lambda}_3=0. \tag{2.19}$$

If none of the λ's or $\widetilde{\lambda}$'s vanish, it is fairly easy to see that no solution exists for *real* four-momenta—that is, for momenta satisfying $\widetilde{\lambda}_a=\pm\lambda_a^*$. Indeed, by contracting (2.19) with various spinors, we find that

$$\lambda_1\widetilde{\lambda}_1+\lambda_2\widetilde{\lambda}_2+\lambda_3\widetilde{\lambda}_3=0 \implies \Big\{\langle 1\,2\rangle[1\,2]=\langle 2\,3\rangle[2\,3]=\langle 3\,1\rangle[3\,1]=0\Big\}. \tag{2.20}$$

But if we consider the λ and $\widetilde{\lambda}$ to be independent (equivalently, consider the momenta to be generally complex), then two solutions exist:

$$\begin{Bmatrix}(A)\colon \lambda_1\propto\lambda_2\propto\lambda_3 \implies \langle 1\,2\rangle=\langle 2\,3\rangle=\langle 3\,1\rangle=0\\ (B)\colon \widetilde{\lambda}_1\propto\widetilde{\lambda}_2\propto\widetilde{\lambda}_3 \implies [1\,2]=[2\,3]=[3\,1]=0\end{Bmatrix}; \tag{2.21}$$

that is, either all square brackets vanish *or* all angle brackets vanish. Because of this, the three-particle scattering amplitude must depend either exclusively on the $\widetilde{\lambda}$'s or exclusively on the λ's—cases (A) and (B), respectively. And because there does not exist any little-group-invariant combination of either set of brackets by themselves, the scaling property of any on-shell function, (2.8), uniquely dictates the kinematical dependence of the S-matrix:

$$\begin{Bmatrix}(A)\colon\ \mathcal{A}_3(h_1,h_2,h_3)\propto[12]^{h_1+h_2-h_3}\,[23]^{h_2+h_3-h_1}\,[31]^{h_3+h_1-h_2}\\ (B)\colon\ \mathcal{A}_3(h_1,h_2,h_3)\propto\langle 12\rangle^{h_3-h_1-h_2}\,\langle 23\rangle^{h_1-h_2-h_3}\,\langle 31\rangle^{h_2-h_3-h_1}\end{Bmatrix}. \tag{2.22}$$

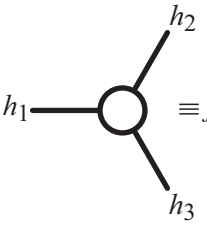
The final constraint is that we must require the amplitude to be smooth in the limit of real momenta—when all brackets vanish, $\langle a\,b\rangle\to\mathcal{O}(\epsilon)$, and $[a\,b]\to\mathcal{O}(\epsilon)$; taking this limit, we see that the expressions (2.22) become

$$\begin{Bmatrix}(A)\colon\ [12]^{h_1+h_2-h_3}\,[23]^{h_2+h_3-h_1}\,[31]^{h_3+h_1-h_2}\xrightarrow[[ab]\to\mathcal{O}(\epsilon)]{}\mathcal{O}\big(\epsilon^{(h_1+h_2+h_3)}\big)\\ (B)\colon\ \langle 12\rangle^{h_3-h_1-h_2}\,\langle 23\rangle^{h_1-h_2-h_3}\,\langle 31\rangle^{h_2-h_3-h_1}\xrightarrow[\langle ab\rangle\to\mathcal{O}(\epsilon)]{}\mathcal{O}\big(\epsilon^{-(h_1+h_2+h_3)}\big)\end{Bmatrix}.$$

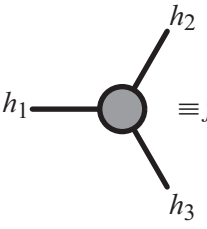
Hence, only case (A) is valid when $\sum_a h_a\geq 0$, and only case (B) is valid when $\sum_a h_a\leq 0$. (When $\sum_a h_a=0$, a combination of both can be used.)

Thus, depending on the helicities of the particles involved, the (analytic continuation of the) three-particle S-matrix for massless particles is uniquely fixed up to an overall coupling constant, f^{q_1,q_2,q_3}, which in general depends on

the non-kinematical quantum numbers q_a labeling distinguishable external states. Using empty and filled vertices to denote the two cases, respectively, we have:



$$\equiv f^{q_1,q_2,q_3}\,[12]^{h_1+h_2-h_3}\,[23]^{h_2+h_3-h_1}\,[31]^{h_3+h_1-h_2}\,\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big); \quad \big(h_1+h_2+h_3\geq 0\big)$$



$$\equiv f^{q_1,q_2,q_3}\,\langle 12\rangle^{h_3-h_1-h_2}\,\langle 23\rangle^{h_1-h_2-h_3}\,\langle 31\rangle^{h_2-h_3-h_1}\,\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big). \quad \big(h_1+h_2+h_3\leq 0\big)$$

(2.23)

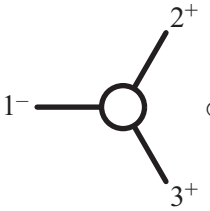
Because the kinematical dependence of the S-matrix is fixed by helicities of the particles involved, a particular quantum field theory is specified merely by listing the observable particles' helicities and the coupling constants f^{q_1,q_2,q_3} for all the allowed three-particle interactions. Although the primary focus of our work involves the *kinematical* structure of on-shell scattering processes—and we will soon begin to leave these non-kinematical coupling constants implicit—a few aspects of these coefficients are worth mentioning here.

First, dimensional analysis shows that the coupling constant f^{q_1,q_2,q_3} has units, carrying the dimensions of mass according to:

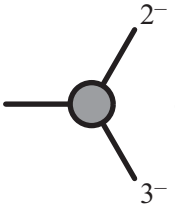
$$\big[f^{q_1,q_2,q_3}\big]=\big[\text{mass}\big]^{1-|h_1+h_2+h_3|}. \tag{2.24}$$

When discussing theories of massless particles, there is a natural preference for those involving only dimensionless coupling constants; and so a special role is played by theories involving only vertices for which $|h_1+h_2+h_3|=1$.

Second, although it may at first appear that any choice of particles and coupling constants could define a suitable quantum field theory, basic considerations of quantum mechanics strongly constrain the possible choices. Consider for example a theory involving only particles with $\sigma_a=1$ (such particles are called "gluons"). Allowing only interactions with dimensionless couplings, there are only two basic interactions (where we denote states with $h_a=\pm 1$ by $+$ and $-$ respectively):



$$\propto \frac{[23]^3}{[12][31]}\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big),$$



$$\propto \frac{\langle 23\rangle^3}{\langle 12\rangle\langle 31\rangle}\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big)$$

where for each, the constant of proportionality is of course f^{q_1,q_2,q_3}, which depends on the non-kinematical quantum numbers (called "colors"), q_a, labeling the various distinguishable particles. As spin-1 particles are bosons, the amplitudes written above must be symmetric under exchanging particles $2 \leftrightarrow 3$; and because the kinematical part is antisymmetric, we must have that $f^{q_1,q_2,q_3} = -f^{q_1,q_3,q_2}$. Extending this argument via crossing symmetry implies that f^{q_1,q_2,q_3} must be fully antisymmetric in the color labels q_a.

In addition to its antisymmetry, there is one more important (and perhaps surprising) constraint from quantum mechanics on the possible coupling constants. Although we will not reproduce the argument here, it was shown in reference [68] that consistency among the factorization channels of the four-particle amplitude requires that the coupling constants satisfy a Jacobi identity:

$$f^{q_1,q_2,\bullet\overline{f}\bullet,q_3,q_4} + f^{q_2,q_3,\bullet\overline{f}\bullet,q_1,q_4} + f^{q_3,q_1,\bullet\overline{f}\bullet,q_2,q_4} = 0. \tag{2.25}$$

(Here, we have used Wick contraction notation for a sum over particle labels.) Because of this, whatever non-kinematical quantum numbers label distinguishable spin-1 particles, they form the adjoint representation of some Lie algebra—thus proving that Yang–Mills theory is the unique theory of interacting spin-1 particles! Similar arguments described in reference [68] show that the interactions of any spin-2 particle (a "graviton") must satisfy the equivalence principle—that is, the mass-scale involved in the graviton self-coupling must be the same as that for its couplings to all other fields. We direct the interested reader to reference [68] for a detailed discussion.

For the rest of this book, we will largely ignore the non-kinematical coupling constants defining the theory as they add nothing to the kinematical structure, and can always be reintroduced *post facto*. But before moving on, there is one last thing worth mentioning: for Yang–Mills (and its supersymmetric extensions), the existence of these couplings provides a completely well-defined limit in which only *planar* scattering processes are relevant. This is because planar on-shell diagrams decorated with f's (when viewed as the structure constants of a Lie algebra via (2.25)) dominate nonplanar diagrams in the limit of large rank. If the Lie algebra were $SU(N)$ (the most physically interesting case) then nonplanar diagrams are suppressed by powers of $1/N$ relative to planar ones (see e.g. reference [69]), so that only planar graphs are relevant when $N \to \infty$. This is good, because it turns out that the set of planar on-shell diagrams are much easier to understand mathematically—being closely linked to what is known as the "positroid stratification" of the Grassmannian. As such, we will mostly focus on planar scattering processes below. But we should reiterate that nonplanar on-shell diagrams are entirely well defined by the rule given in equation (2.16).

2.4 On-shell supersymmetry and (maximally) supersymmetric theories

In theories with supersymmetry, particles with different helicities are unified into groups called super-multiplets—where all the component particles in a given multiplet have identical couplings to other fields in the theory. These particles are connected to one another by $\mathcal{N}$ supersymmetry generators Q_I^α and $\widetilde{Q}_I^{\dot\alpha}$, with $SU(\mathcal{N})$ "R-charge" indices $I = 1,\ldots,\mathcal{N}$, defined so that (see e.g. [65, 70–73])

$$\left\{Q_I^\alpha, \widetilde{Q}_J^{\dot\alpha}\right\} = 2\delta_{IJ} P^{\alpha\dot\alpha}, \tag{2.26}$$

where $P^{\alpha\dot\alpha}$ is the standard momentum operator. Notice that this identification makes clear an analogy between the supersymmetry generators and the spinor-helicity operators; indeed, when acting on external states, they act according to

$$Q_I^\alpha |a\rangle^{h_a} = \lambda_a^\alpha |a\rangle_I^{h_a+1/2} \quad \text{and} \quad \widetilde{Q}_I^{\dot\alpha} |a\rangle^{h_a} = \widetilde{\lambda}_a^{\dot\alpha} |a\rangle_I^{h_a-1/2}. \tag{2.27}$$

Introducing a conjugate spinor $\widetilde{\lambda}_{\overline{a}}$ for $\widetilde{\lambda}_a$ such that $[\overline{a}\,a] = 1$, we can absorb the spinor-index of $\widetilde{Q}_I^{\dot\alpha}$ by introducing $\widetilde{Q}_I \equiv \epsilon_{\dot\alpha\dot\beta}\widetilde{Q}_I^{\dot\alpha}\widetilde{\lambda}_{\overline{a}}^{\dot\beta}$. With this notational simplification, we can generate a super-multiplet by successively acting on the highest-helicity state (for which $Q_I^\alpha |a\rangle^{+\sigma_a} = 0$) with $\widetilde{Q}_I$ as follows:

state	degeneracy
$\|a\rangle^{+\sigma_a} \equiv \|a\rangle^{\sigma_a}$	1
$\widetilde{Q}_I\|a\rangle^{+\sigma_a} \equiv \|a\rangle_I^{\sigma_a-1/2}$	$\mathcal{N}$
$\widetilde{Q}_I\widetilde{Q}_J\|a\rangle^{+\sigma_a} \equiv \|a\rangle_{IJ}^{\sigma_a-1}$	$\mathcal{N}(\mathcal{N}-1)/2$
$\vdots \quad \vdots$	$\vdots$
$\widetilde{Q}_I\cdots\widetilde{Q}_K\|a\rangle^{+\sigma_a} \equiv \|a\rangle_{I\cdots K}^{\sigma_a-\mathcal{N}/2}$	1

(2.28)

In general, it would be necessary to introduce a CPT-conjugate super-multiplet by acting successively on the state with helicity $(\sigma_a - \mathcal{N}/2)$ with the operator Q. But this is not necessary when $\mathcal{N} = 4$ *and* the highest-helicity state has spin $\sigma_a = 1$, because then the super-multiplet is CPT self-conjugate. This theory is called *maximally supersymmetric Yang–Mills*, or simply '$\mathcal{N} = 4$' for short, and will be the primary example studied in this book.

When $\mathcal{N} = 4$, we may collect together the various helicity states into an eigenstate of the $\widetilde{Q}$-operators, called a *Grassmann coherent state* [65]:

$$|a\rangle \equiv e^{\widetilde{Q}_I^{\dot\alpha}\widetilde{\lambda}_{\overline{a},\dot\alpha}\widetilde{\eta}_a^I}|a\rangle^{+1}, \tag{2.29}$$

where we have introduced anti-commuting ('fermionic') parameters $\widetilde{\eta}_a^I$ to label the various helicity states in the multiplet. Expanding (2.29), we see that:

$$|a\rangle \equiv |a\rangle^{+1} + \widetilde{\eta}^I|a\rangle_I^{+1/2} + \frac{1}{2!}\widetilde{\eta}_a^I\widetilde{\eta}_a^J|a\rangle_{IJ}^0 + \frac{1}{3!}\widetilde{\eta}_a^I\widetilde{\eta}_a^J\widetilde{\eta}_a^K|a\rangle_{IJK}^{-1/2} + \frac{1}{4!}\widetilde{\eta}_a^I\widetilde{\eta}_a^J\widetilde{\eta}_a^K\widetilde{\eta}_a^L|a\rangle_{IJKL}^{-1}.$$

Notice that because the $\widetilde{\eta}$'s are anti-commuting, this multiplet contains one pair of particles with $h_a = \pm 1$, four pairs with $h_a = \pm\frac{1}{2}$, and six particles with $h_a = 0$.

Thus, scattering amplitudes in $\mathcal{N} = 4$ are functions not of the helicities h_a, but of the anti-commuting ("fermionic") parameters $\widetilde{\eta}_a^I$ that label the component fields of each external state. Under the action of the little group, $\widetilde{\eta}_a$ transforms like $\widetilde{\lambda}_a$, so that an external super-multiplet has uniform weight, and any on-shell superfunction will transform according to:

$$f(\cdots, t_a\lambda_a, t_a^{-1}\widetilde{\lambda}_a, t_a^{-1}\widetilde{\eta}_a, \cdots) = t_a^{-2} f(\cdots, \lambda_a, \widetilde{\lambda}_a, \widetilde{\eta}_a, \cdots). \tag{2.30}$$

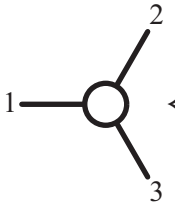
Just as the external states are polynomials in the $\widetilde{\eta}$'s, any on-shell superfunction will be a polynomial in the $\widetilde{\eta}$'s—component functions being given as the coefficient of certain combinations of $\widetilde{\eta}$'s according to (2.29). Because any superfunction $f(\lambda, \widetilde{\lambda}, \widetilde{\eta})$ must be invariant under the $SU(4)$ R-symmetry, it can be decomposed into terms of uniform degree in the $\widetilde{\eta}$'s:

$$f(\lambda, \widetilde{\lambda}, \widetilde{\eta}) \equiv \sum_k f^{(k)}(\lambda, \widetilde{\lambda}, \widetilde{\eta}), \tag{2.31}$$

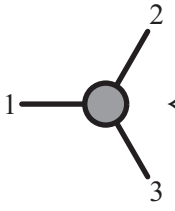
where $f^{(k)}$ is of degree $4k$ in the $\widetilde{\eta}$'s. A superfunction $f^{(k)}$ of degree $4k$ will have among its components those functions involving k 'negative helicity' gluons (particles with $h_a = -1$), and $(n-k)$ 'positive helicity' gluons (particles with $h_a = +1$).

(Sometimes, we will describe a superfunction $f^{(k)}(\lambda, \widetilde{\lambda}, \widetilde{\eta})$ as "N$^{(k-2)}$MHV," where 'MHV' stands for 'maximal helicity-violating' and 'N' denotes 'next to;' this terminology follows from the fact that (for $n > 3$ particles) super-momentum conservation (see below) implies that superfunctions with $k < 2$ must vanish, suggesting that superfunctions with $k = 2$ be called 'MHV.')

Because all the helicity states in $\mathcal{N} = 4$ are unified into a single super-multiplet, the superamplitude will contain among its components all possible helicity amplitudes with fixed $\sum_a h_a$. Thus, there is a unique three-particle S-matrix of each type described in section 2.3. In notation we will clarify momentarily, these are:



$$\Leftrightarrow \mathcal{A}_3^{(1)} \equiv \frac{\delta^{1\times 4}\left(\widetilde{\lambda}^{\perp}\cdot\widetilde{\eta}\right)}{[1\,2][2\,3][3\,1]}\delta^{2\times 2}\left(\lambda\cdot\widetilde{\lambda}\right),\quad \widetilde{\lambda}^{\perp} \equiv \left([2\,3][3\,1][1\,2]\right);$$



$$\Leftrightarrow \mathcal{A}_3^{(2)} \equiv \frac{\delta^{2\times 4}\left(\lambda\cdot\widetilde{\eta}\right)}{\langle 1\,2\rangle\langle 2\,3\rangle\langle 3\,1\rangle}\delta^{2\times 2}\left(\lambda\cdot\widetilde{\lambda}\right). \tag{2.32}$$

Notice that the $\widetilde{\eta}$-dependence of each amplitude is organized into a simple anti-commuting ("fermionic") δ-function of the form $\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)$, where C is some $(k\times n)$ matrix of ordinary functions: in the case of $\mathcal{A}_3^{(2)}$, the matrix C is simply λ; while for $\mathcal{A}_3^{(1)}$ it is the (1×3) matrix $\widetilde{\lambda}^{\perp}$ given above (the notation following from the fact that $\widetilde{\lambda}^{\perp}\cdot\widetilde{\lambda}=[2\,3]\widetilde{\lambda}_1+[3\,1]\widetilde{\lambda}_2+[1\,2]\widetilde{\lambda}_3=0$ for any three generic 2-spinors $\widetilde{\lambda}_a$).

To be (perhaps overly) explicit, given some $(k\times n)$ matrix C of ordinary functions, $\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)$ simply denotes the *product* of all its $4k$ components. Making use of the antisymmetry of the $\widetilde{\eta}$'s, the expansion becomes a collection of $\binom{n}{k}^4$ component functions, each involving a choice of four $(k\times k)$ minors of C:

$$\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)\equiv\prod_{I=1}^{4}\Big(\sum_{1\le a_1<\cdots<a_k\le n}(a_1\cdots a_k)\widetilde{\eta}^I_{a_1}\cdots\widetilde{\eta}^I_{a_k}\Big),\tag{2.33}$$

with $(a_1\cdots a_k)\equiv\det\{c_{a_1},\ldots,c_{a_k}\}$, and c_a a column vector of $C\equiv(c_1\cdots c_n)$. This follows from the definition of fermionic integration,

$$\int d\widetilde{\eta}^I_a\;\delta\big(c_a\,\widetilde{\eta}^I_a\big)=\int d\widetilde{\eta}^I_a\;\big(c_a\,\widetilde{\eta}^I_a\big)=c_a,\tag{2.34}$$

together with the anti-commutation properties of the $\widetilde{\eta}$'s.

Just to be completely clear about this notation, let us use it to see how a multitude of component, helicity amplitudes are encoded in each of the three-point amplitudes given in (2.32). Using the (1×3) matrix $\widetilde{\lambda}^{\perp}$ defined in (2.32), we see that the δ-function in $\mathcal{A}_3^{(1)}$ can be expanded according to the rule in (2.33) as:

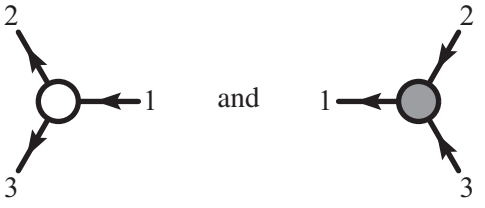
$$\begin{aligned}\delta^{1\times 4}\big(\widetilde{\lambda}^{\perp}\cdot\widetilde{\eta}\big)=\;&\big([2\,3]\widetilde{\eta}^1_1\big)\big([2\,3]\widetilde{\eta}^2_1\big)\big([2\,3]\widetilde{\eta}^3_1\big)\big([2\,3]\widetilde{\eta}^4_1\big)+\ldots\;\text{(3 terms)}\\&+\big([2\,3]\widetilde{\eta}^1_1\big)\big([2\,3]\widetilde{\eta}^2_1\big)\big([2\,3]\widetilde{\eta}^3_1\big)\big([3\,1]\widetilde{\eta}^4_2\big)+\ldots\;\text{(24 terms)}\\&+\big([2\,3]\widetilde{\eta}^1_1\big)\big([2\,3]\widetilde{\eta}^2_1\big)\big([3\,1]\widetilde{\eta}^3_2\big)\big([1\,2]\widetilde{\eta}^4_3\big)+\ldots\;\text{(36 terms)}\\&+\big([2\,3]\widetilde{\eta}^1_1\big)\big([2\,3]\widetilde{\eta}^2_1\big)\big([3\,1]\widetilde{\eta}^3_2\big)\big([3\,1]\widetilde{\eta}^4_2\big)+\ldots\;\text{(18 terms).}\end{aligned}\tag{2.35}$$

Thus, the superamplitude $\mathcal{A}_3^{(1)}$ can be thought of as simply denoting a collection of $81(=\binom{3}{1}^4)$ component helicity amplitudes indexed by the $\widetilde{\eta}$'s:

$$\begin{aligned}\mathcal{A}_3^{(1)}=\;&\big(\widetilde{\eta}^1_1\big)\big(\widetilde{\eta}^2_1\big)\big(\widetilde{\eta}^3_1\big)\big(\widetilde{\eta}^4_1\big)\mathcal{A}\big(|1\rangle^{-1}_{1234},|2\rangle^{+1},\;|3\rangle^{+1}\big)\;+\ldots\;\text{(3 terms)}\\&+\big(\widetilde{\eta}^1_1\big)\big(\widetilde{\eta}^2_1\big)\big(\widetilde{\eta}^3_1\big)\big(\widetilde{\eta}^4_2\big)\mathcal{A}\big(|1\rangle^{-1/2}_{123},|2\rangle^{+1/2}_{4},|3\rangle^{+1}\big)\;+\ldots\;\text{(24 terms)}\\&+\big(\widetilde{\eta}^1_1\big)\big(\widetilde{\eta}^2_1\big)\big(\widetilde{\eta}^3_2\big)\big(\widetilde{\eta}^4_3\big)\mathcal{A}\big(|1\rangle^{0}_{12},\;|2\rangle^{+1/2}_{3},|3\rangle^{+1/2}_{4}\big)+\ldots\;\text{(36 terms)}\\&+\big(\widetilde{\eta}^1_1\big)\big(\widetilde{\eta}^2_1\big)\big(\widetilde{\eta}^3_2\big)\big(\widetilde{\eta}^4_2\big)\mathcal{A}\big(|1\rangle^{0}_{12},\;|2\rangle^{0}_{34},\;|3\rangle^{+1}\big)\;+\ldots\;\text{(18 terms).}\end{aligned}\tag{2.36}$$

The inherent *simplicity* of $\mathcal{N}=4$ is reflected in the fact that there are *only two* fundamental three-particle (super)amplitudes that define the theory—without the need to specify the helicities of the particular states involved at any vertex.

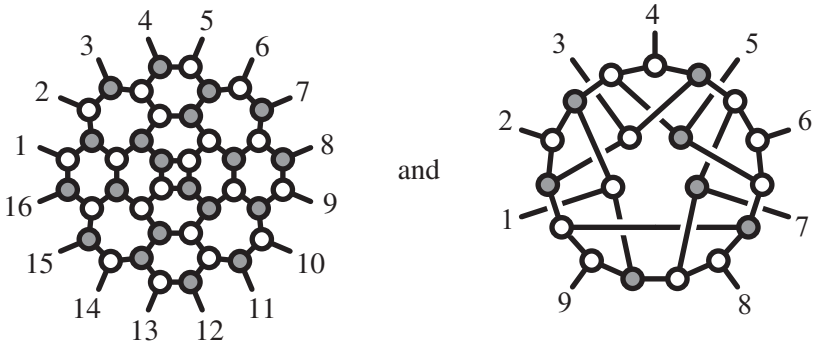
In contrast, for a theory with $\mathcal{N} < 4$ (including non-supersymmetric ('pure') Yang–Mills), there are effectively *three* distinct three-particle vertices of each type, depending on the states involved. If we denote the flow of helicity by directing the edges at each vertex, the fundamental amplitudes defining pure Yang–Mills, for example, would be the pair of amplitudes



(2.37)

together with their cyclically rotated manifestations. We can of course amalgamate such vertices into on-shell diagrams defined for less supersymmetric theories, but the resulting on-shell diagrams would depend on the orientations given to the edges—which is more information than that required for an on-shell diagram of $\mathcal{N} = 4$. Nevertheless, we will see in Chapter 15 that the additional structure required to describe on-shell functions in a theory with $\mathcal{N} < 4$ supersymmetries amounts to an overall prefactor (depending on the orientations of the edges) relative to what would be computed for an undirected on-shell diagrams relevant to $\mathcal{N} = 4$.

Because the fundamental three-particle amplitudes required to describe $\mathcal{N} = 4$ are *undirected*, the set of on-shell diagrams obtained through their amalgamations are correspondingly simpler. By following the amalgamation procedure described in section 2.2, the simple three-particle amplitudes can be used to generate on-shell diagrams and functions as complex as the following examples:



(2.38)

Here, we have drawn both planar and nonplanar examples to emphasize that nothing about the amalgamation procedure involves planarity. This being said, however, our present understanding of planar on-shell diagrams is much more comprehensive and our (mathematical) tools of analysis are much more refined; therefore, we will focus primarily on planar diagrams in this work.

When computing diagrams in $\mathcal{N}=4$, the general rule for amalgamating amplitudes at vertices, (2.16), is greatly simplified because the sum over each internal particle's helicity can be re-expressed as a Grassmann integral over its $\widetilde{\eta}$:

$$\prod_{i\in I}\left(\int\frac{d^2\lambda_i\,d^2\widetilde{\lambda}_i}{\mathrm{vol}(GL(1))}d^4\widetilde{\eta}_i\right)\prod_v \mathcal{A}_v. \tag{2.39}$$

We can also think of this as a generalization of on-shell phase space to an on-shell *super*-phase space according to:

$$\frac{d^2\lambda_a\,d^2\widetilde{\lambda}_a}{\mathrm{vol}(GL(1))}d^4\widetilde{\eta}_a. \tag{2.40}$$

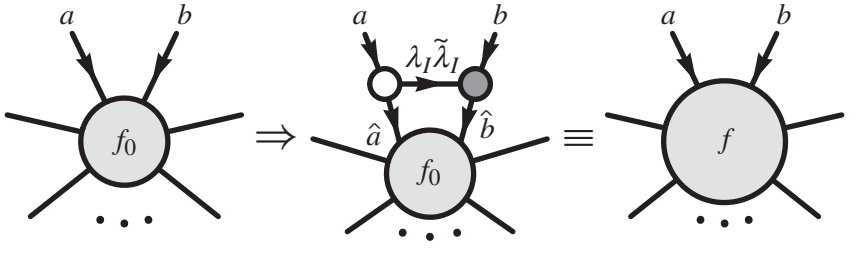
Similarly, ordinary momentum conservation is upgraded to *super*-momentum conservation, which becomes the constraint:

$$\delta^{2\times 4}\big(\lambda\cdot\widetilde{\eta}\big)\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big). \tag{2.41}$$

Thus, super-momentum conservation would seem to require that $k\geq 2$ for any non-vanishing on-shell diagram. The only exception is for $n=3$ particles, in which case λ can be replaced by the (1×3) matrix $\widetilde{\lambda}^{\perp}$ defined in (2.32); this makes sense, as the solution to three-particle momentum conservation relevant for the $k=1$ vertex is when $\langle 12\rangle=\langle 23\rangle=\langle 31\rangle=0$, which implies that λ is effectively a rank-1 matrix of spinors (effectively, $\lambda\subset\widetilde{\lambda}^{\perp}$). It is for this reason that diagrams for which $k=2$ are somewhat special, and are occasionally referred to as "maximal helicity violating" ('MHV') as mentioned above.

2.5 Building up diagrams with "BCFW-bridges"

One particularly simple way of building up more complicated on-shell diagrams from simpler ones will play an important role in our story. Starting from any on-shell diagram f_0, we can pick two external lines, and attach a "BCFW-bridge" to make a new diagram f as follows:



(2.42)

(The arrows here merely indicate the direction of momentum-flow.)

Let us determine explicitly how the on-shell function f is related to that of f_0. Using the general rule for translating a diagram into an on-shell function, (2.39),

and observing that the bridge adds three new internal lines and two new vertices, we see that f is given by

$$f = \int \prod_{i\in\{\widehat{a},\widehat{b},I\}} \frac{d^2\lambda_i\, d^2\widetilde{\lambda}_i}{\mathrm{vol}(GL(1))} d^4\widetilde{\eta}_i\, \mathcal{A}_3^{(1)}(a,I,\widehat{a})\mathcal{A}_3^{(2)}(b,\widehat{b},I) f_0(\ldots,\widehat{a},\widehat{b},\ldots). \tag{2.43}$$

Now, the two three-particle vertices include the momentum-conserving δ-functions,

$$\delta^{2\times 2}\big(\lambda_a\widetilde{\lambda}_a - \lambda_I\widetilde{\lambda}_I - \lambda_{\widehat{a}}\widetilde{\lambda}_{\widehat{a}}\big)\delta^{2\times 2}\big(\lambda_b\widetilde{\lambda}_b - \lambda_{\widehat{b}}\widetilde{\lambda}_{\widehat{b}} + \lambda_I\widetilde{\lambda}_I\big), \tag{2.44}$$

and the fermionic δ-functions,

$$\delta^{1\times 4}\big([I\widehat{a}]\widetilde{\eta}_a - [\widehat{a}a]\widetilde{\eta}_I - [aI]\widetilde{\eta}_{\widehat{a}}\big)\delta^{2\times 4}\big(\lambda_b\widetilde{\eta}_b - \lambda_{\widehat{b}}\widetilde{\eta}_{\widehat{b}} + \lambda_I\widetilde{\eta}_I\big). \tag{2.45}$$

Therefore, the eight ordinary δ-functions and the twelve fermionic δ-functions can be used to fully localize *all but one* of the 3×3 internal phase space integrals, and *all* of the 3×4 fermionic integrals. Upon localizing as many phase space integrals as possible, with one remaining degree of freedom associated with λ_I, we find:

$$\left\{\begin{array}{l}\lambda_{\widehat{a}} = \lambda_a \\ \widetilde{\lambda}_{\widehat{a}} = \widetilde{\lambda}_a - \alpha\,\widetilde{\lambda}_b \\ \widetilde{\eta}_{\widehat{a}} = \widetilde{\eta}_a - \alpha\,\widetilde{\eta}_b\end{array}\right\},\quad \left\{\begin{array}{l}\lambda_I = \alpha\,\lambda_a \\ \widetilde{\lambda}_I = \widetilde{\lambda}_b \\ \widetilde{\eta}_I = \widetilde{\eta}_b\end{array}\right\},\quad \left\{\begin{array}{l}\lambda_{\widehat{b}} = \lambda_b + \alpha\,\lambda_a \\ \widetilde{\lambda}_{\widehat{b}} = \widetilde{\lambda}_b \\ \widetilde{\eta}_{\widehat{b}} = \widetilde{\eta}_b\end{array}\right\}. \tag{2.46}$$

Notice that solving the δ-functions contributes an overall Jacobian factor $\alpha\langle ab\rangle[ab]$ to the denominator of the result. Using the isolated solutions (2.46), and taking the appropriate $\widetilde{\eta}$-components, we find that the three-point amplitudes become

$$\frac{[aI]^4}{[aI][I\widehat{a}][\widehat{a}a]}\frac{\langle\widehat{bI}\rangle^4}{\langle b\widehat{b}\rangle\langle\widehat{bI}\rangle\langle Ib\rangle} = \frac{[ab]^4}{[ab][ab](\alpha[ab])}\frac{(\alpha\langle ab\rangle)^4}{(\alpha\langle ab\rangle)(\alpha\langle ab\rangle)(\alpha\langle ab\rangle)} = \langle ab\rangle[ab],$$

which, when combined with the Jacobian from the δ-functions, is simply $1/\alpha$!

Thus, attaching a BCFW-bridge adds one new degree of freedom, α, to an on-shell function f_0, giving rise to a new on-shell function f according to:

$$\begin{aligned} f(\ldots,\lambda_a,\widetilde{\lambda}_a,\widetilde{\eta}_a,\lambda_b,\widetilde{\lambda}_b,\widetilde{\eta}_b,\ldots) &= \frac{d\alpha}{\alpha} f_0(\ldots,\lambda_{\widehat{a}},\widetilde{\lambda}_{\widehat{a}},\widetilde{\eta}_{\widehat{a}},\lambda_{\widehat{b}},\widetilde{\lambda}_{\widehat{b}},\widetilde{\eta}_{\widehat{b}},\ldots), \\ &= \frac{d\alpha}{\alpha} f_0(\ldots,\lambda_a,\widetilde{\lambda}_a - \alpha\,\widetilde{\lambda}_b,\widetilde{\eta}_a - \alpha\,\widetilde{\eta}_b,\lambda_b + \alpha\,\lambda_a,\widetilde{\lambda}_b,\widetilde{\eta}_b,\ldots). \end{aligned} \tag{2.47}$$

In particular, notice that adding a BCFW-bridge introduces one new (simple) *pole*—the residue about which results in the diagram without the bridge. So BCFW-bridges act in some ways like Feynman propagators—although instead of introducing a quadratic pole involving *off-shell* degrees of freedom, the bridge introduces only a simple pole, arising from purely *on-shell* degrees of freedom.

By successively attaching BCFW-bridges to a small set of 'simple' diagrams, a very complex collection of diagrams can be produced (both planar and nonplanar

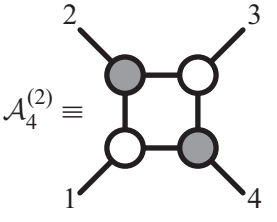
alike). Indeed, as we will soon understand, it turns out that *all* (physically relevant) on-shell diagrams can be constructed in this way.

2.6 On-shell recursion for all-loop amplitudes

While on-shell diagrams are interesting in their own right, for planar $\mathcal{N}=4$, we will see that they are of much more than purely formal interest. Scattering amplitudes at all loop orders can be *directly represented* and computed as on-shell scattering processes. This is quite remarkable, considering the ubiquity of "off-shell" data in the more familiar Feynman expansion.

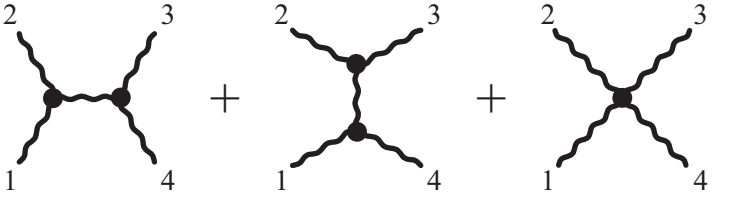
Of course by now we have become accustomed to the idea that amplitudes can be 'determined' using on-shell data—as evidenced, for instance, by the BCFW recursion relations at tree [74, 75] and loop levels [14] (see also [76–79]). But our statement goes beyond this: the claim is not just that an off-shell object such as "the loop *integral*" can be *determined* using only on-shell information, but rather that loop *integrands* can be *directly represented* by fully on-shell objects.

Before discussing loops, let us look at some examples of "tree-level" amplitudes. Recall from [80] that the four-particle tree-amplitude $\mathcal{A}_4^{(2)}$ can be represented by a single on-shell diagram—its "BCFW representation" (an example of (2.18)):

$\mathcal{A}_4^{(2)} \equiv$ 

(2.48)

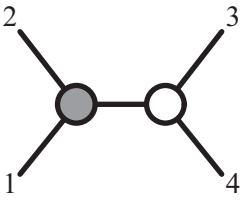
This is very far from the representation given by the Feynman expansion, which would have expressed (2.48) as the sum of three Feynman diagrams,



(2.49)

the first two of which involve *off-shell* gluon exchange. (The terms "tree-amplitude" and "loop-amplitude" are artifacts of such Feynman-diagrammatic expansions.) Another striking difference is that, despite the fact that we're discussing a *tree*-amplitude, the on-shell diagram (2.48) looks like a loop! To emphasize this distinction, consider a (possibly more familiar) "tree-like" on-shell diagram

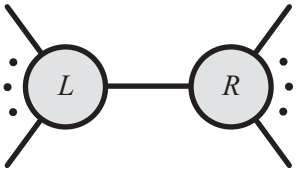
such as:



(2.50)

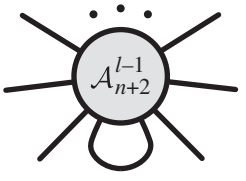
Since the internal line must be on-shell, the diagram imposes a δ-function constraint, $\delta((p_1+p_2)^2)$, on the external momenta; and so, (2.50) corresponds to a singularity—a *factorization channel*. The extra leg in (2.48) that makes the "loop" allows for a non-vanishing result for *generic* (on-shell, momentum-conserving) external momenta. It is interesting to note that we can interpret (2.48) as having been obtained by attaching a "BCFW-bridge" to *any* of the factorization channels of the four-particle amplitude—such as that of (2.50). This makes it possible for the single diagram, (2.48), to *simultaneously* exhibit *all* the physical factorization channels.

This simple example illustrates the fundamental physical idea behind the BCFW description of an amplitude—not just at tree level, but at all loop orders: any amplitude can be *fully* reconstructed from the knowledge of its singularities; and the singularities of an amplitude are determined entirely by on-shell data. At tree level, the singularities are simply the familiar factorization channels,



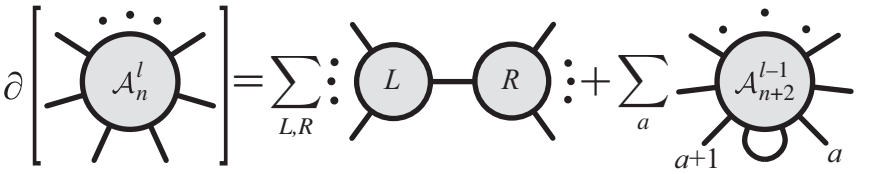
(2.51)

where the left- and right-hand sides are both on-shell scattering amplitudes. At loop level, all the singularities of the integrand can be understood as factorizations like that of (2.51), or those for which an *internal* particle is put on-shell; at least for $\mathcal{N}=4$ in the planar limit, these singularities are given by the "forward-limit" [81] of an on-shell amplitude with one fewer loop and two extra particles, where any two adjacent particles have equal and opposite momenta, denoted:



(2.52)

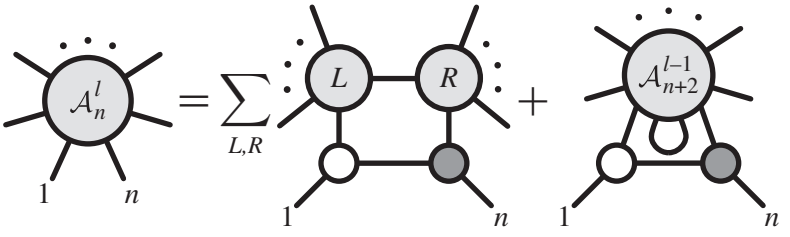
Combining these two terms, the singularities of the *full amplitude* are [14]:



(2.53)

Here we have suggestively used the symbol "∂" to signify "singularity of". Of course, the symbol ∂ is often used to denote "boundary" or "derivative"; we will soon see that all of these senses are appropriate.

Equation (2.53) can be understood as defining a "differential equation" for scattering amplitudes; and it turns out to be possible to 'integrate' it directly. This is precisely what is accomplished by the BCFW recursion relations [14]. For planar $\mathcal{N}=4$, the all-loop BCFW recursion relations represent the l-loop amplitude, $\mathcal{A}_n^l$, in terms of on-shell diagrams according to:



(2.54)

The structure of this solution will be discussed in much greater detail in Chapter 17. For instance, notice that this presentation makes only *some* of the factorization channels and forward-limits manifest, and seems to break the cyclic symmetry of the amplitude by singling out legs $(1\,n)$. In other words, working intrinsically with on-shell diagrams, it is not obvious that the singularities of the sum (2.54) includes *all* the required singularities of the amplitude. Of course Feynman diagrams *do* make it manifest that such an object exists; but it would be nice to understand this more directly, without recourse to the usual formalism of field theory. We will show how this works in section 17.1, demonstrating that (2.54) has all the necessary singularities purely from within the framework of on-shell diagrams.

The seed of loop integrands in the recursion relation are the "forward-limit" terms—as the three-point amplitudes are fixed by Poincaré invariance to all loop orders. Each loop is accompanied by four integration variables: three of these are given by the phase space of the forward-limit momentum $\lambda_{AB}\widetilde{\lambda}_{AB}$ (from merging legs 'A' and 'B'), and the BCFW deformation parameter α is the fourth. Of course, all the objects appearing in these expressions are completely on-shell, and so do not seem to contain anything that looks like the conventional "$\int d^4\ell$" to which we are

accustomed (where ℓ is the momentum of a generally off-shell, *virtual* particle). However, it is easy to convert the parameters of the on-shell forward-limit to the more familiar one via the identification

$$\ell \equiv \lambda_{AB}\widetilde{\lambda}_{AB} + \alpha\, \lambda_1 \widetilde{\lambda}_n \text{ with } d^4\ell = \frac{d^2\lambda_{AB} d^2\widetilde{\lambda}_{AB}}{\text{vol}(GL(1))} d\alpha\, \langle 1\, \lambda_{AB}\rangle [n\, \widetilde{\lambda}_{AB}]. \tag{2.55}$$

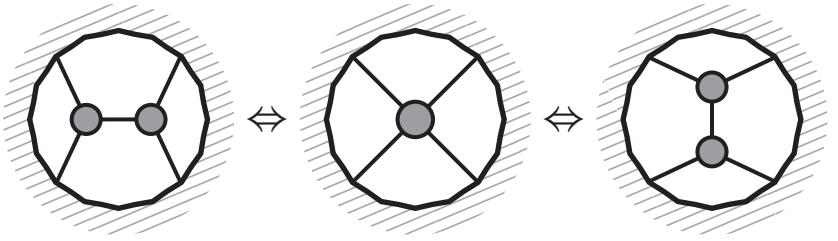
At l loops, the recursion relations (2.54) produce a $4l$-form on internal phase space, and we can identify the $4l$ integration variables with loop momenta at each order via (2.55). Integrating these on-shell forms over a contour that restricts each loop-momentum to be *real* (i.e. in $\mathbb{R}^{3,1}$) generates the ultimate physical amplitude.

Thus, as advertised, on-shell diagrams are of much more than mere academic interest: they *fully* determine the amplitude in planar $\mathcal{N} = 4$ to *all loop-orders*.

2.7 Physical equivalences among on-shell diagrams

We have seen that on-shell diagrams are objects of fundamental importance to the physics of scattering amplitudes. It is therefore natural to try and compute the on-shell forms associated with on-shell diagrams more explicitly, and better understand their structure. At first sight, the class of on-shell diagrams may look as complicated as Feynman diagrams. For instance, even for a fixed number of external particles, there are obviously an infinite number of such diagrams (by continuously adding BCFW-bridges, for example). As we will see however, at least for $\mathcal{N} = 4$ in the planar limit, this complexity is entirely illusory. The reason is that apparently very different graphs actually give rise to exactly the same on-shell differential forms—differing only by a change of variables.

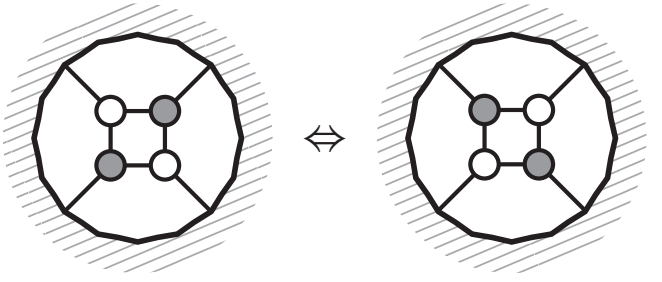
The first instance of this phenomenon is extremely simple and trivial. Consider an analog of the "factorization channel" diagram (2.50), but connecting two black vertices. Because these vertices require that all the $\widetilde{\lambda}$'s be parallel, it makes no physical difference how they are connected. And so, on-shell diagrams related by



(2.56)

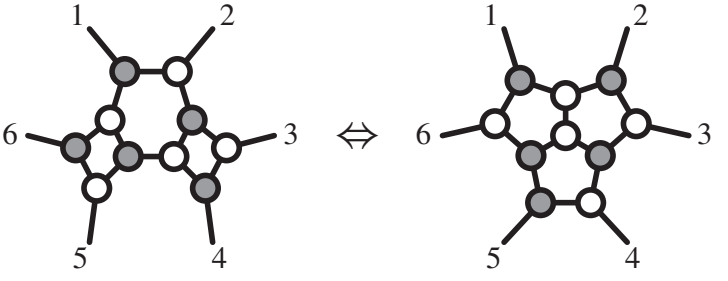
represent the same on-shell form. Thus, we can collapse and re-expand any chain of connected black vertices in any way we like; the same is obviously true for white vertices. Because of this, for some purposes it may be useful to define composite black-and-white vertices with any number of legs. By grouping black and white vertices together in this way, on-shell diagrams can always be made *bipartite*—with (internal) edges only connecting white with black vertices. We will, however, preferentially draw trivalent diagrams because of the fundamental role played by the three-particle amplitudes.

There is also a more interesting equivalence between on-shell diagrams that will play an important role in our story. We can see this already in the BCFW representation of the 4-particle amplitude given above, (2.48). The *picture* is obviously not cyclically invariant—as a rotation would exchange its black and white vertices. But the four-particle *amplitude* of course *is* cyclically invariant; and so there is another generator of equivalences among on-shell diagrams, the "square move" [82]:



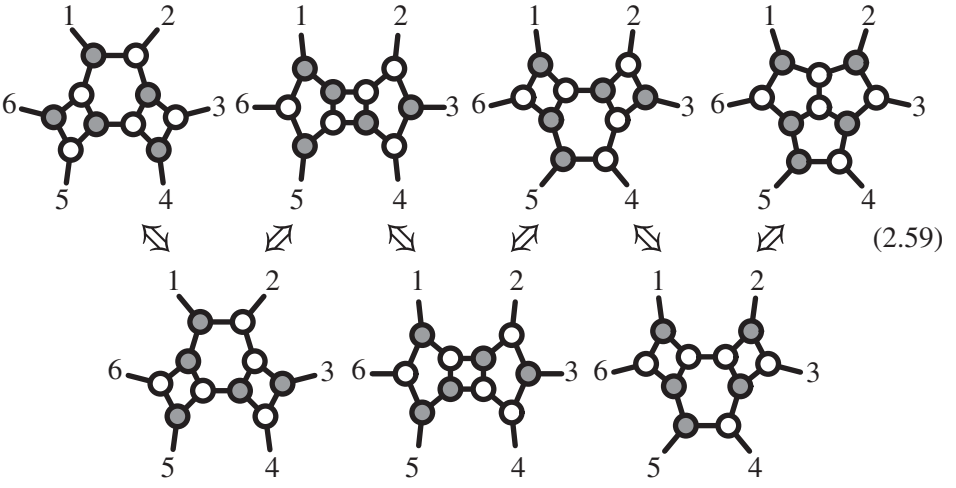
(2.57)

The merge and square moves can be used to show the physical equivalence of many seemingly different on-shell diagrams. For instance, the following two diagrams generate physically equivalent on-shell forms:



(2.58)

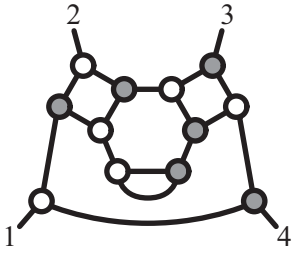
We can see this by explicitly constructing the chain of moves that brings one graph into the other:



(2.59)

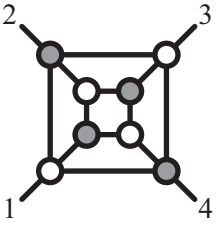
Here, each step down involves one or more square moves, and each step up involves one or more mergers.

To give another example, the on-shell diagram representing the one-loop four-particle amplitude—obtained directly from BCFW recursion, (2.54)—is given by:



(2.60)

Using a series of mergers and square moves, it can be brought to the beautifully symmetric, bipartite form:



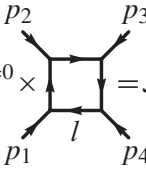
(2.61)

These forms are completely equivalent, but suggest very different physical interpretations. The first, (2.60), clearly exposes its origin as a forward-limit—arising through the gluing of two of the external legs of a six-particle tree-amplitude. The second, (2.61), does not look like this at all; instead, it appears as four BCFW-bridges attached to an internal square—which is of course the four-particle tree-amplitude. Thus, in this picture, we can think of the 1-loop amplitude as an integral over a four-parameter deformation of the tree-amplitude!

This is more than mere amusement. It immediately tells us that, with an appropriate choice of variables representing the BCFW-shifts, the one-loop amplitude can be represented in a remarkably simple form:

$$\mathcal{A}_4^{l=1} \propto \mathcal{A}_4^{l=0} \times \int \frac{d\alpha_1}{\alpha_1}\frac{d\alpha_2}{\alpha_2}\frac{d\alpha_3}{\alpha_3}\frac{d\alpha_4}{\alpha_4} = \mathcal{A}_4^{l=0} \times \int d\log(\alpha_1)\cdots d\log(\alpha_4)\,. \quad (2.62)$$

Of course, this does not look anything like the more familiar expression [83]

$\mathcal{A}_4^{l=1} \propto \mathcal{A}_4^{l=0} \times$  $= \mathcal{A}_4^{l=0} \times \int \dfrac{d^4\ell\ (p_1+p_2)^2(p_1+p_3)^2}{\ell^2(\ell+p_1)^2(\ell+p_1+p_2)^2(\ell-p_4)^2}\,.$ (2.63)

In this representation, it is not at all obvious that there is any change of variables that reduces the integrand to the "$d$log"-form of (2.62). However, following the rule for identifying off-shell loop momenta in terms of on-shell data, (2.55), we may easily identify the map that takes us from the ℓ of (2.63) to the α_i of (2.62):

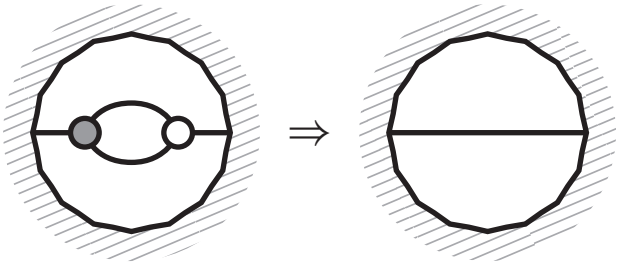
$$\begin{aligned}&\frac{d^4\ell\ (p_1+p_2)^2(p_1+p_3)^2}{\ell^2(\ell+p_1)^2(\ell+p_1+p_2)^2(\ell-p_4)^2} \\ &= d\log\left(\frac{\ell^2}{(\ell-\ell^*)^2}\right) d\log\left(\frac{(\ell+p_1)^2}{(\ell-\ell^*)^2}\right) d\log\left(\frac{(\ell+p_1+p_2)^2}{(\ell-\ell^*)^2}\right) d\log\left(\frac{(\ell-p_4)^2}{(\ell-\ell^*)^2}\right),\end{aligned} \quad (2.64)$$

where ℓ^* is either of the two points null-separated from all four external momenta. This expression will be derived in detail in section 17.3.

As we will see, the existence of a "$d$log" representation for loop integrands is a completely general feature of all amplitudes at all loop-orders. But the possibility of such a form even existing was never anticipated from the more traditional formulations of field theory. Indeed, even for the simple example of the four-particle one-loop amplitude, the existence of a change of variables converting $d^4\ell$ to four $d$log's went unnoticed for decades. We will see that these "$d$log"-forms follow directly from the on-shell diagram description of scattering amplitudes generated by the BCFW recursion relations, (2.54). Beyond their elegance, these $d$log-forms suggest a completely new way of carrying out loop integrations, and hints at an underlying, "motivic" structure of the final results, which will likely be the subject of much further research.

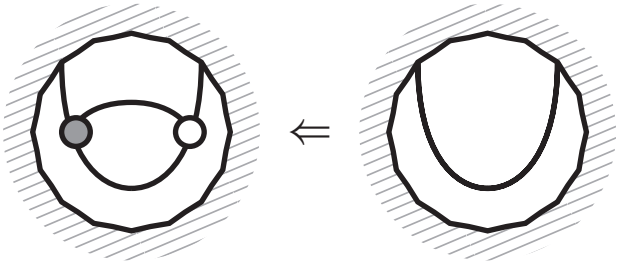
The equivalence of on-shell diagrams related by mergers and square moves clearly represents a major simplification in the structure on-shell diagrams; but these alone cannot reduce the seemingly infinite complexities of graphs with arbitrary numbers of 'loops' (faces) as neither of these operations affects the number of faces of a graph. However, using mergers and square moves, it may

be possible to represent an on-shell diagram in a way that exposes a "bubble" on an internal line. As one might expect, there is a sense in which such diagrams can be *reduced*:



(2.65)

Of course this cannot literally be true: there is one more integration variable in the diagram with the bubble than the one without it. To see this, notice that the diagram with the bubble can be obtained by attaching a (trivial) BCFW-bridge:



(2.66)

(The general discussion of attaching BCFW-bridges in section 2.5 hence provides a rigorous definition of an on-shell function for an on-shell diagram with a bubble.) Therefore, what "reduction" actually means is that there is a concrete *and simple* change of variables for which this extra degree of freedom, say α, factors out of the on-shell form simply as an overall $d\log(\alpha)$—which, upon taking the residue on a contour around $\alpha = 0$, yields the on-shell form for the reduced diagram.

Before completing our discussion, it is worth mentioning that there are other—somewhat trivial—operations on diagrams that leave the corresponding on-shell form invariant; these include, adding or deleting a bivalent vertex (of either color) along a line and exchanging the colors involved in a bubble such as that in (2.65).

It turns out that using mergers, square moves, *and* bubble-deletion, all planar on-shell diagrams involving n external particles can be reduced to a *finite* number of diagrams. This shows that the essential content of on-shell diagrams are encapsulated by the finite list of *reduced* objects. And as we will see, the extra, "irrelevant" variables associated with bubble-deletion also have a purpose in life: they represent loop integration variables, for example.

Reduced diagrams are still not unique of course: they can be transmuted into each other using mergers and square moves. Because the same on-shell function

can be represented by many different on-shell diagrams, it is natural to ask for some *invariant* way to characterize them. For instance, if we are given two complicated on-shell diagrams such as those of (2.58), how can we decide whether they can be morphed into each other using mergers and square moves? The answer to this question turns out to be simple and striking: the invariant data associated with a reduced on-shell diagram is encoded by a *permutation* among the external particle labels! We will describe this connection in detail in the next section.

It is amazing that a connection between scattering amplitudes in $(3+1)$ dimensions and combinatorics exists at all, let alone that it will play a central role in the story. This is the tip of an iceberg of remarkable connections between on-shell diagrams and rich mathematical structures only recently explored in the literature. We will spend much of the rest of this book outlining these connections in greater detail. But recall that this is not the first time that scattering theory has been related to permutations in an important way: a classic example of such a connection is for integrable theories in $(1+1)$ dimensions. In addition to providing us with some historical context, revisiting this story will give us an interesting perspective on recent developments.

3

Permutations and scattering amplitudes

3.1 Combinatorial descriptions of scattering processes

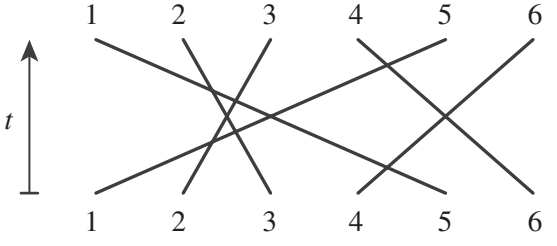
To a physicist, scattering is perhaps the most fundamental physical process; but scattering amplitudes are rather sophisticated functions of the helicities and momenta of the external particles. If we strip away all of this data, all that would be left would be the arbitrary labels identifying the particles involved, which we will denote simply by $\{1,\ldots,n\}$. The simplest kind of "interaction" that could be associated with just this data would be a *permutation*; because of the central role played by permutations in combinatorics, we might fancifully say that permutations are combinatorial analogs of the S-matrix.

At first sight, it certainly seems as if a "combinatorial S-matrix" would be far too simple an object to capture anything remotely resembling the richness of physical scattering amplitudes. However, we will soon discover that this is not the case: on-shell diagrams are fully determined by permutations, and hence the entire S-matrix of $\mathcal{N}=4$ can be described combinatorially!

Recall that something very much like this happens for integrable theories in $(1+1)$ dimensions [84, 85]. Consider for instance the permutation given by

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 3 & 2 & 6 & 1 & 4 \end{pmatrix}. \tag{3.1}$$

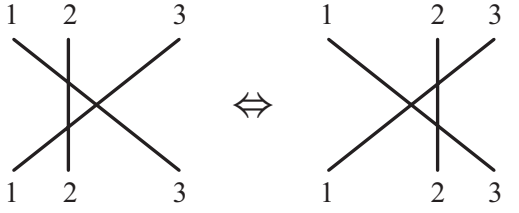
Its relationship to physics can be seen by representing it graphically as:



(3.2)

This can be thought of as a space-time picture for a scattering process in $(1+1)$ dimensions, where time flows upwards. First, particles 4 and 5 scatter, then 1 and

2, then 2 and 3, and so on. The time-ordering of these scatterings corresponds to one way of representing the permutation as a product of adjacent transpositions. Of course, this decomposition is not unique: there are many ways of drawing the same picture with different time-orderings for the various $2 \to 2$ processes. In a general theory with only 4-point interactions, the amplitude for different orderings would be different, and therefore the amplitude for the scattering process would not be completely determined by the permutation alone. For the amplitude to depend *only on the permutation* and *nothing else*, the $2 \to 2$ amplitudes must satisfy the famous Yang–Baxter relation [84, 85]:



(3.3)

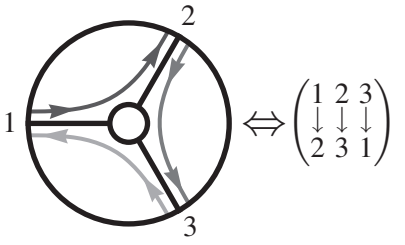
It is natural to ask whether such a picture can be generalized to more realistic theories in higher dimensions. This seems impossible at first sight, since the pictures drawn above only make physical sense in $(1+1)$ dimensions (not only because they are drawn on a plane). The fact that particles can only move in one spatial dimension is what makes it possible to describe all interactions as a sequence of local $2 \to 2$ scattering processes. Also important is the absence of any particle creation or destruction, allowing us to label the final states by the same labels as the initial states. Neither of these features hold for the higher-dimensional theories in which we are primarily interested: for planar $\mathcal{N} = 4$, particle creation and destruction play a fundamental role; and the most primitive processes are not $2 \to 2$ amplitudes, but rather the 3-particle amplitudes discussed above, (2.32).

An important starting point for describing higher-dimensional scattering processes is to forgo the traditional meaning of the "S-matrix"—an operator that maps initial states to final states. Rather, we find it much more convenient to treat all the external particles on equal footing, using crossing symmetry to formulate the S-matrix as a process for which *all* the external particles are taken to be *incoming*.

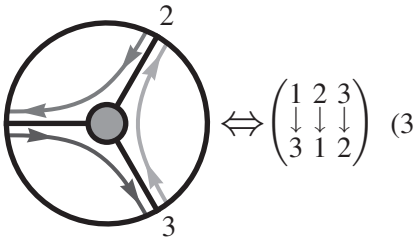
One lesson we can take from $(1+1)$ dimensions is that any connection between scattering and permutations *must* involve *on-shell* processes. In $(3+1)$ dimensions, this leads us to trivalent, on-shell diagrams with black and white vertices discussed in the previous chapter. And so we are led to try and associate a permutation with these diagrams. As it turns out, just such a connection exists between planar, bi-colored graphs (called "plabic" graphs) and permutations, which has recently been studied in the mathematics literature [39] (see also [42]).

Let's jump in and describe how it works. The way to read off a permutation from an on-shell graph is as follows. For each external leg a (with clockwise ordering),

follow the graph inward from a, turning left at each white vertex, and turning right at each black vertex; this "left-right path" will terminate at some external leg, denoted $\sigma(a)$. For example, the three-particle building blocks of $\mathcal{N}=4$, (2.32), are associated with permutations in the following way:

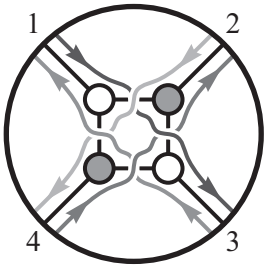


$$\Leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1 \end{pmatrix} \quad \text{and} \quad$$



$$\Leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 1 & 2 \end{pmatrix} \tag{3.4}$$

Of course, this works equally well for more complex on-shell graphs; for example, the graph giving the four-particle tree-amplitude, (2.48), is associated with the following permutation:



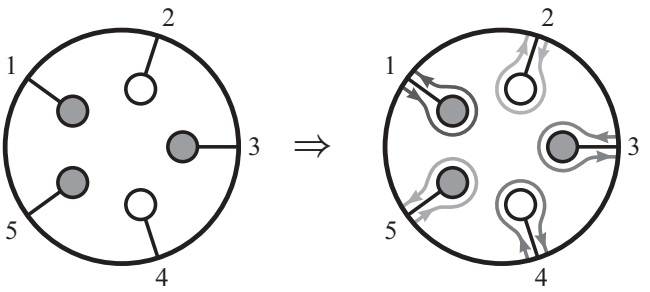
$$\Leftrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 4 & 1 & 2 \end{pmatrix} \tag{3.5}$$

It is very easy to see that such "left-right paths" allow us to define a permutation for any planar graph constructed with black and white vertices (not only those that are trivalent). Starting from any external leg of such a graph, this path will always lead back out to the boundary; and because any path can be trivially reversed (by exchanging the roles of black and white), it is clear that every external leg is the terminus of some such path. And so, the left-right paths do indeed define a *permutation* among the external legs.

Actually, left-right paths associate each graph with a slight generalization of an ordinary permutation known as a *decorated* permutation—a generalization that allows for two types of fixed-points. By convention, we always consider a left-right path to permute each label 'to its right'—in other words, we think of the paths as being associated with a map $\sigma : \{1,\ldots,n\} \mapsto \{1,\ldots,2n\}$ such that $a \leq \sigma(a) \leq a+n$; taking $\sigma(a) \mod n$ would be an ordinary permutation. The two types of fixed points correspond to the cases of $\sigma(a) = a$ or $\sigma(a) = a+n$. For the sake of simplicity, for the rest of this book we will refer to these *decorated* permutations simply as 'permutations' and denote them by "$\{\sigma(1),\ldots,\sigma(n)\}$".

To emphasize how "decorated" permutations differ from ordinary permutations, notice that we may differentiate among 2^n decorations of the identity permutation.

These are relevant for on-shell diagrams such as

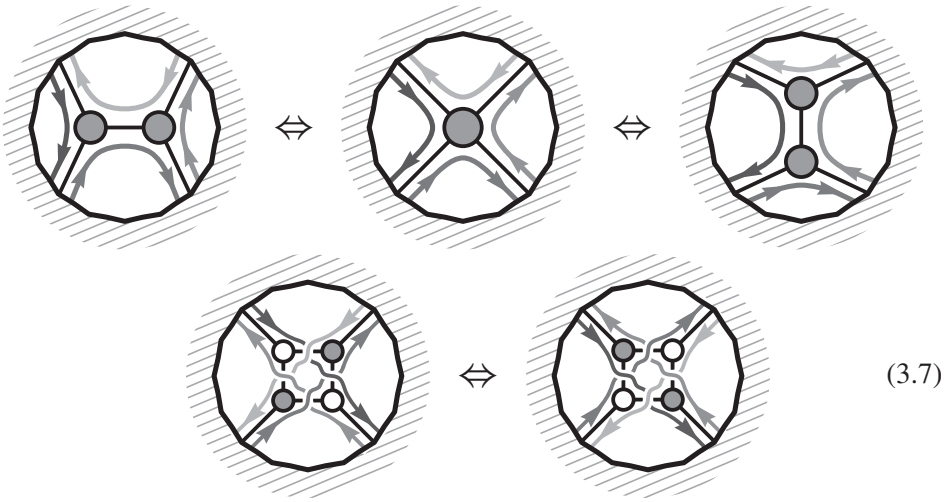


(3.6)

which would be labeled by a 'permutation' $\{1,7,3,9,5\}$. Although such empty graphs are themselves of little direct relevance to physics, they will play an important role in the general toolbox—as we will see in the following section.

Associated with any permutation is a number, k, which is the number of legs $a \in \{1,\ldots,n\}$ that are mapped 'beyond n' by σ—that is, for which $\sigma(a) > n$. This number is also given by the mean value of $\sigma(a) - a$: $k \equiv \frac{1}{n}\sum_a(\sigma(a) - a)$. To see this, notice that while this mean for any *ordinary* permutation always vanishes, our requirement that $a \leq \sigma(a) \leq a + n$ means that σ must be shifted by n relative to an ordinary permutation for some k elements. For example, both the 4-point graph, (3.5), and the 5-particle graph, (3.6), have $k = 2$.

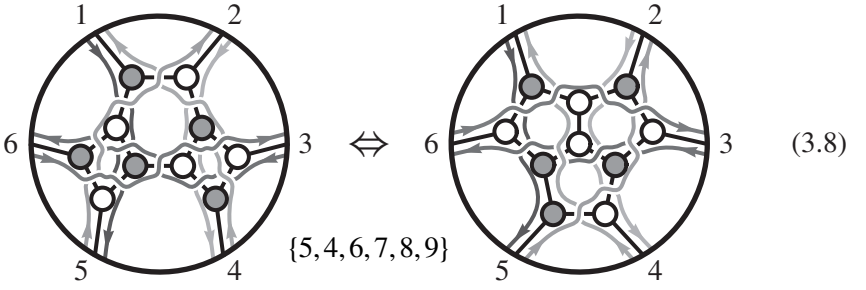
The reason why the permutations associated with on-shell graphs are so important is that in many cases they *invariantly* encode the physical information about the graph and the on-shell form associated with it. Recall that graphs related by mergers, (4.63), or square moves, (2.57), represent the same physical, on-shell function. These operations also leave permutations invariant:



(3.7)

Bubble-deletion, however, *does* change the permutation associated with an on-shell diagram; it also changes the number of faces. But by deleting bubbles, any graph can be 'reduced'—and any two *reduced* graphs labeled by the same

permutation always represent the same physical form. More explicitly, all physical information in reduced graphs is captured by the corresponding permutation. To see a simple example of this, recall the pair of inequivalent graphs given in (2.58) that were related by a rather long sequence of mergers and square moves; it is much easier to test the equivalence of the permutations that label them:



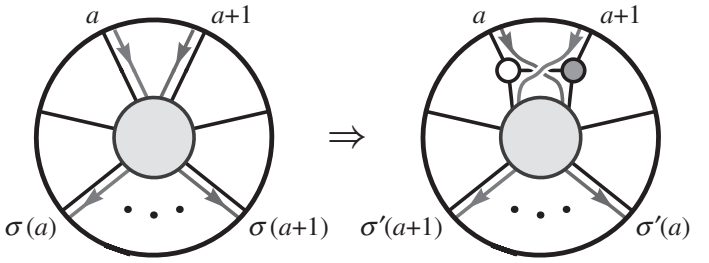
(3.8)

We should note in passing that there is something very special about $\mathcal{N}=4$ and integrability that allows us to *fully* characterize on-shell diagrams in this way. Just as the Yang–Baxter relation (3.3) was the prerequisite for $(1+1)$-dimensional theories to be entirely combinatorial, it is the square move (2.57) that does this for $\mathcal{N}=4$: recall that in a non-supersymmetric theory, all 3-particle vertices would need to be dressed by the helicities of the particles involved—such as in (2.25); this dressing represents *extra* data that must be supplied in order to specify a physical process, and this data is *not* left invariant under square moves. That being said, however, the purely combinatorial story of $\mathcal{N}=4$ will play a central role even for non-supersymmetric theories. This will be described in Chapter 15.

3.2 The BCFW-bridge construction of representative graphs

We have seen that every on-shell graph is associated with a permutation; quite beautifully, the converse is also true: *all* permutations can be represented by an on-shell graph. A constructive procedure for building a representative graph for any permutation was described in [39] (and in somewhat different terms by D. Thurston in [42]). Here, we will describe a different method—motivated by simple physical and combinatorial considerations, and by analogy with physics in $(1+1)$ dimensions where graphs are constructed out of simple, adjacent transpositions. Of course, in $(3+1)$ dimensions, there is no space-time evolution analogue of successive $2\to 2$ scattering; and so we must find some way to 'build up' on-shell objects directly from the "vacuum" (a trivial permutation).

The key is understanding what an adjacent transposition means in terms of on-shell graphs. The answer is extremely simple: an adjacent transposition is just exactly the addition of a BCFW-bridge (see section 2.5):



(3.9)

Notice that any number of 'hanging legs'—those which map to themselves under σ—can be inserted between a and "$a+1$" without consequence; and so, we will consider any transposition $(a\,c)$ to be "adjacent" so long as for all b between a and c, $\sigma(b) = b \bmod n$. (Although the bridge drawn in (3.9) will be sufficient for most applications, the oppositely colored bridge—where black and white vertices are exchanged—could also be used; the difference being that such a bridge would transpose the *pre*-images of a and $a+1$ under σ instead of their images.)

Because adjacent transpositions simply correspond to adding BCFW-bridges, any decomposition of a permutation σ into a sequence of such transpositions acting on a trivial permutation can be read as *instructions* for building-up a *representative* on-shell graph for σ by successively adding BCFW-bridges to an empty graph like that of (3.6). Of course, adding a BCFW-bridge may potentially give us a reducible on-shell diagram. However, it turns out that when adding a bridge to a reduced graph, so long as $\sigma(a+1) < \sigma(a)$—that is, the paths are arranged as drawn in (3.9)—then the resulting graph is guaranteed to be reduced. We will not prove this statement now, but its proof will become trivial after the discussions in Chapter 5.

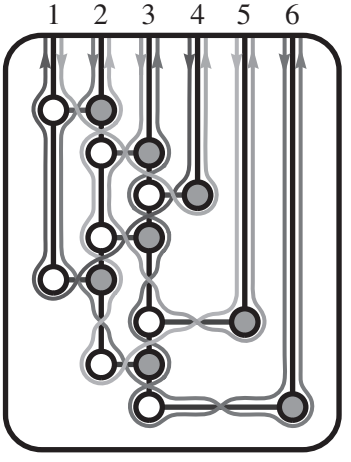
And so, when breaking down a permutation into adjacent transpositions, we want to find pairs $(a\,c)$ with $a < c$ (separated only by legs b self-identified under σ) such that $\sigma(a) < \sigma(c)$; then when we decompose σ as $(a\,c) \circ \sigma'$ with $\{\sigma(a), \sigma(c)\} = \{\sigma'(c), \sigma'(a)\}$, adding a BCFW-bridge to a *reduced* on-shell diagram labeled by σ' will result in a *reduced* on-shell diagram labeled by σ. Of course, there are *many* ways of decomposing a permutation σ into such a chain of adjacent transpositions, and *any* such decomposition will result in a representative, reduced graph whose left-right permutation is σ. But for the sake of concreteness, let us describe one very specific (in fact canonical) procedure to decompose any permutation—one which will turn out to have rather special properties discussed in section 6.4.

BCFW-bridge decomposition: Starting with any permutation σ, if σ is not a decoration of the identity, then decompose σ as $(a\,c) \circ \sigma'$, where $1 \le a < c \le n$ is the *lexicographically first* pair separated only by legs b that are self-identified under σ and for which $\sigma(a) < \sigma(c)$; repeat until σ' is a decoration of the identity.

To illustrate this procedure, let's see how it generates a representative reduced on-shell diagram that is labeled by the permutation $\{4,6,5,7,8,9\}$:

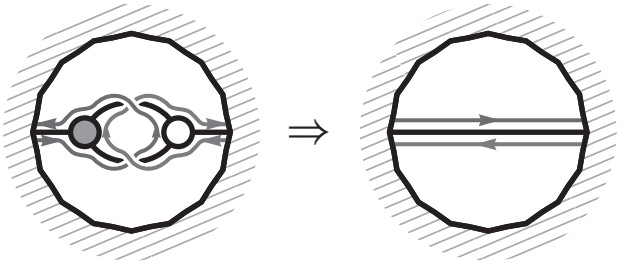
τ	1 ↓	2 ↓	3 ↓	4 ↓	5 ↓	6 ↓
	4	6	5	7	8	9
(1 2)	6	4	5	7	8	9
(2 3)	6	5	4	7	8	9
(3 4)	6	5	7	4	8	9
(2 3)	6	7	5	4	8	9
(1 2)	7	6	5	4	8	9
(3 5)	7	6	8	4	5	9
(2 3)	7	8	6	4	5	9
(3 6)	7	8	9	4	5	6

$\Leftrightarrow$



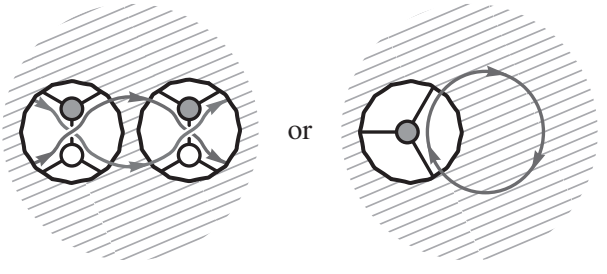
(3.10)

This procedure provides us with a combinatorial test of a graph's reducibility: because the BCFW-bridge construction always produces a *reduced* representative graph for any permutation, and each step in the construction adds one face to the graph as it is built, a graph is reduced if and only if the number of its faces minus one is equal to the number of steps in the BCFW-bridge decomposition of the permutation that labels it. If not, then the graph is reducible, and has some number of faces that can be deleted by bubble-reduction:



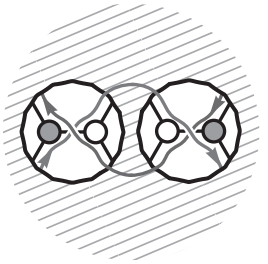
(3.11)

A more intrinsic way to identify a reducible graph is if any pair of left-right paths $a \to \sigma(a)$ and $b \to \sigma(b)$ cross each other along more than one edge in the graph in the manner known as a "bad double-crossing", or if there is any purely internal path:



(3.12)

A bad double-crossing is distinguished from those double-crossings of the form



(3.13)

which do *not* indicate that a graph is reducible.

We thus have a complete dictionary between (reduced) on-shell graphs and permutations. As we will discuss in Chapter 14, this new picture actually contains the $(1+1)$-dimensional story as a special case. Another closely related special case is relevant for describing on-shell diagrams (and all-loop amplitudes) of the ABJM theory in $(2+1)$ dimensions! But let us now move beyond the purely combinatorial aspects of the story, and turn towards actually *computing* on-shell diagrams. This will lead us to uncover beautiful structures in algebraic geometry also described by decorated permutations, ultimately connecting on-shell graphs to the "positive" Grassmannian—a recurring theme in our studies.

4

From on-shell diagrams to the Grassmannian

In this chapter we will show that the computation of on-shell diagrams is most efficiently and transparently carried out by associating each diagram with an auxiliary structure: a matrix C representing a subspace of the Grassmannian $G(k,n)$. But let us begin by reviewing some elementary properties about Grassmannian manifolds in general, and describe the first appearance of these spaces in the story of scattering amplitudes, as they arise in the description of external kinematical data.

4.1 The Grassmannian of k-planes in n dimensions, $G(k,n)$

The Grassmannian $G(k,n)$ is the space of k-dimensional planes passing through the origin in an n-dimensional space (see e.g. [66]). We can specify a k-plane in n dimensions by giving k vectors $c^\alpha \in \mathbb{C}^n$, whose span defines the plane. We can assemble these vectors into a $(k\times n)$ matrix C, whose components are c^α_a for $\alpha = 1,\ldots,k$ and $a = 1,\ldots,n$.

Under $GL(k)$-transformations, $C \mapsto \Lambda \cdot C$—with $\Lambda \in GL(k)$—the row vectors will change, but the plane spanned by them is obviously unchanged. Thus, the Grassmannian $G(k,n)$ can be thought of as the space of $(k\times n)$ matrices modulo this $GL(k)$ "gauge" redundancy. From this, we see that the dimension of $G(k,n)$ is $k\times n - k^2 = k(n-k)$. In practice, we can "gauge-fix" the $GL(k)$ redundancy by choosing any k of the columns of the matrix to form the $(k\times k)$ identity matrix. For instance, we can represent a generic point in $G(2,5)$ in the following gauge-fixed form:

$$C = \begin{pmatrix} 1 & 0 & c_3^1 & c_4^1 & c_5^1 \\ 0 & 1 & c_3^2 & c_4^2 & c_5^2 \end{pmatrix}. \tag{4.1}$$

This coordinate chart does not cover the *entire* Grassmannian—though of course the collection of all $\binom{n}{k}$ such charts would obviously suffice.

The $GL(k)$-invariant information associated with C is easily specified. First, notice that the only $SL(k)$-invariants of $C \in G(k,n)$ are the minors constructed out of the *columns* of C,

$$(a_1 \cdots a_k) \equiv \det\{c_{a_1}, \ldots, c_{a_k}\}. \tag{4.2}$$

$GL(k)$-invariants are then simply ratios of these:

$$\frac{(a_1 \cdots a_k)}{(b_1 \cdots b_k)}. \tag{4.3}$$

While the (ratios of) minors are $GL(k)$-invariant, the number of these, $\binom{n}{k}$, is much greater than the dimensionality of the Grassmannian, $\dim(G(k,n)) = k(n-k)$, and so the minors represent a highly redundant set of data to describe C. The identities among minors arise from the simple fact that any k-vector can be expanded in a basis of any k linearly independent k-vectors—a statement that is equivalent to the identity known as *Cramer's rule*:

$$c_{a_1}(a_2 \cdots a_{k+1}) - c_{a_2}(a_1 a_3 \cdots a_{k+1}) + \cdots + (-1)^k c_{a_{k+1}}(a_1 \cdots a_k) = 0, \tag{4.4}$$

for any $c_a \in \mathbb{C}^k$. Contracting each of the vectors in (4.4) with another set of vectors $c_{b_1}, \ldots, c_{b_{k-1}}$ generates the identities known as the Plücker relations,

$$(b_1 \cdots b_{k-1} a_1)(a_2 \cdots a_{k+1}) + \cdots + (-1)^{k-1}(b_1 \cdots b_{k-1} a_{k+1})(a_1 \cdots a_k) = 0. \tag{4.5}$$

Associated with any k-plane C is a natural $(n-k)$-plane denoted $C^\perp$, the "orthogonal complement" of C, which is defined by

$$C^\perp \cdot C = 0. \tag{4.6}$$

Therefore, there is a natural isomorphism between $G(k,n)$ and $G(n-k,n)$, which is reflected in the invariance of $\dim(G(k,n)) = k(n-k)$ under the exchange $k \leftrightarrow (n-k)$. The minors of $C^\perp$ are fully determined by the minors of C in the obvious way: for any complementary sets $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_{n-k}\}$ (whose union is $\{1, \ldots, n\}$), we have

$$(a_1 \cdots a_k)|_C = \pm (b_1 \cdots b_{n-k})|_{C^\perp}. \tag{4.7}$$

To be completely explicit, suppose we represent C in a gauge where columns c_A with $A \equiv \{a^1, \ldots, a^k\}$ are taken as the identity; then the $n-k$ columns of C in the complementary set $B \equiv A^c$, c_b for $b \in B$—whose components we write as c_b^A—encode the $k(n-k)$ degrees of freedom of C; then the matrix $C^\perp$ has

components

$$\left(c^{\perp}\right)_A^b = -c_b^A. \tag{4.8}$$

For example, the plane $C^{\perp} \in G(3,5)$ orthogonal to $C \in G(2,5)$ given in (4.1) is:

$$C^{\perp} \equiv \begin{pmatrix} (c^{\perp})_1^3 & (c^{\perp})_2^3 & 1 & 0 & 0 \\ (c^{\perp})_1^4 & (c^{\perp})_2^4 & 0 & 1 & 0 \\ (c^{\perp})_1^5 & (c^{\perp})_2^5 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -c_3^1 & -c_3^2 & 1 & 0 & 0 \\ -c_4^1 & -c_4^2 & 0 & 1 & 0 \\ -c_5^1 & -c_5^2 & 0 & 0 & 1 \end{pmatrix}. \tag{4.9}$$

Finally, we will eventually be talking about a certain top-dimensional differential form on the Grassmannian, so it is useful to discuss what general forms on the Grassmannian look like in the coordinates c_a^{α}. Consider first the familiar example of a form on the projective space $G(1,2)$. We can think of this as a (1×2) matrix $C=(c_1\, c_2)$, modulo the $GL(1)$-action of $C \to tC$. Any top-form can be written as

$$\Omega = \frac{d^2C}{\mathrm{vol}(GL(1))} \frac{1}{f(C)}, \tag{4.10}$$

where $f(C)$ must have homogeneity $(+2)$ under rescaling C; that is, $f(tC) = t^2 f(C)$. In practice, "modding-out" by the $GL(1)$-action is trivial: one can simply gauge-fix the $GL(1)$ so that, say, $C \mapsto C^* = (1\, c_2)$; and then $\Omega = dc_2/f(C^*)$. We can also say this more invariantly, by writing,

$$\Omega = \langle CdC \rangle \frac{1}{f(C)}. \tag{4.11}$$

The generalization of this simple case to an arbitrary Grassmannian is straightforward. We can write

$$\Omega = \frac{d^{k\times n}C}{\mathrm{vol}(GL(k))} \frac{1}{f(C)}, \tag{4.12}$$

where $GL(k)$-invariance implies, in particular, that $f(C)$ must be a function of the minors of C with homogeneity under rescaling

$$f(tC) = t^{k\times n} f(C). \tag{4.13}$$

In the coordinate chart where we gauge-fix k of the columns to the identity as above, then $\Omega = d^{k\times(n-k)} c_a^{\alpha}/f(C)$. Said more invariantly, we have

$$\Omega = \langle c^1 \cdots c^k (d^{(n-k)} c^1) \rangle \cdots \langle c^1 \cdots c^k (d^{(n-k)} c^k) \rangle \frac{1}{f(C)}, \tag{4.14}$$

where c^{α} is a row vector of $C \equiv c_a^{\alpha}$ and, e.g.,

$$\langle c^1 \cdots c^k (d^{(n-k)} c^1) \rangle \equiv \epsilon^{a_1 a_2 \ldots a_n} c_{a_1}^1 \cdots c_{a_k}^k \left(dc_{a_{k+1}}^1 \wedge \cdots \wedge dc_{a_n}^1 \right). \tag{4.15}$$

4.2 Grassmannian description of kinematical data—λ and $\tilde{\lambda}$

In a moment, we will establish a very direct connection between on-shell diagrams and the Grassmannian; but let us first pause to point out an even more basic way in which the Grassmannian makes an appearance in scattering amplitudes—in the very way we encode external kinematical data. We normally think of this data as simply being specified by n 2-component spinors λ_a^α and $\widetilde{\lambda}_a^{\dot{\alpha}}$; but of course we may also think of this data as given by a pair of $(2\times n)$ matrices—which we denote collectively by λ and $\widetilde{\lambda}$. For example, the λ's are naturally associated with the $(2\times n)$ matrix,

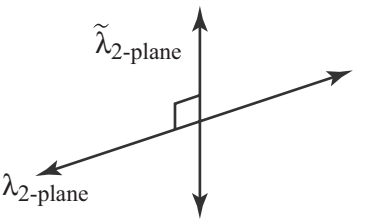
$$\lambda \equiv \begin{pmatrix} \lambda_1^1 & \lambda_2^1 & \cdots & \lambda_n^1 \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n \end{pmatrix}. \tag{4.16}$$

Instead of focusing on the *columns* of the matrix λ, let us think about it as two row vectors. Each of these is a vector in an n-dimensional space. Under Lorentz transformations, these two vectors change, but since Lorentz transformations act on the λ's by $SL(2)$-transformations on their α indices, the two new vectors will simply be a linear combination of the original ones. Therefore, while the vectors themselves change, the *plane* that is spanned by them is invariant under Lorentz transformations. Quite beautifully then, the Lorentz-invariant information encoded by the λ's is really just this 2-*plane* in n dimensions—an element of $G(2,n)$ as realized in [15]. The same is obviously true for the $\widetilde{\lambda}$'s. Of course, the Lorentz group is only the $SL(2)$ part of $GL(2)$ and on-shell forms do transform under "global" little-group transformations that correspond to the $GL(1)$ subgroup of $GL(2)$.

In terms of spinor helicity variables, momentum conservation is simply

$$\sum_a \lambda_a^\alpha \widetilde{\lambda}_a^{\dot{\alpha}} = 0, \tag{4.17}$$

which has the geometric interpretation that the plane λ is *orthogonal to* the plane $\widetilde{\lambda}$ [15]:



(4.18)

This geometric understanding of momentum-conservation also nicely explains the unique nature of its application to the case of three particles: two 2-planes in three dimensions *cannot* be orthogonal in general. A solution, therefore, would be for one of the planes to actually be a 1-plane in disguise. For example, suppose that we have three generic $\widetilde{\lambda}$'s. Momentum conservation requires that $\lambda \subset \widetilde{\lambda}^\perp$, but $\widetilde{\lambda}^\perp$

is a 1-plane! A $GL(1)$-representative of $\widetilde{\lambda}^{\perp}$ is given by

$$\widetilde{\lambda}^{\perp} \equiv \big([23]\ [31]\ [12]\big), \tag{4.19}$$

for which $\widetilde{\lambda}^{\perp}\cdot\widetilde{\lambda}=0$ follows as a trivial instance of Cramer's rule, (4.4):

$$\widetilde{\lambda}^{\perp}\cdot\widetilde{\lambda} = [23]\widetilde{\lambda}_1 + [31]\widetilde{\lambda}_2 + [12]\widetilde{\lambda}_3 = 0. \tag{4.20}$$

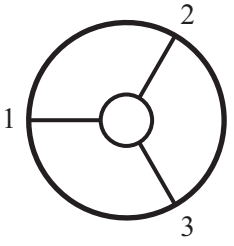
Because this is the *unique* plane orthogonal to $\widetilde{\lambda}$, momentum conservation requires that the λ-plane be spanned by it. In particular, this means that all the λ's must be proportional: in a Lorentz frame where $\lambda_1 = \binom{[23]}{0}$, we have

$$\lambda \equiv \big(\lambda_1\ \lambda_2\ \lambda_3\big) = \begin{pmatrix} [23] & [31] & [12] \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.21}$$

4.3 Grassmannian representation of on-shell diagrams

Let us begin to more explicitly calculate the differential form associated with a given on-shell diagram. We use the momentum-conserving δ-functions at the vertices to localize as many of the internal momenta as we can. This looks highly nontrivial because momentum conservation is a quadratic constraint on the $\lambda, \widetilde{\lambda}$ in general. But a moment's reflection suggests that the situation may be easier to understand. We know that for 3-particle amplitudes, momentum conservation implies a very simple geometric situation—where either the λ's or the $\widetilde{\lambda}$'s are forced to be parallel to each other. However, our representation of the 3-particle amplitude, simple and elegant though it is, does not make this simple fact manifest. This motivates us to try to express the 3-particle amplitude in a slightly different form—one which makes the geometry of the λ's and $\widetilde{\lambda}$'s in each case as transparent as possible.

Let's start with the $\mathcal{A}_3^{(1)}$ vertex:



$$\Leftrightarrow\ \mathcal{A}_3^{(1)} = \frac{\delta^{1\times 4}([23]\widetilde{\eta}_1 + [31]\widetilde{\eta}_2 + [12]\widetilde{\eta}_3)}{[12][23][31]}\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big). \tag{4.22}$$

Notice that the coefficients of the $\widetilde{\eta}$'s are the same as the factors that appear in the denominator of $\mathcal{A}_3^{(1)}$, and coincide with the 1-plane $\widetilde{\lambda}^{\perp}$ orthogonal to $\widetilde{\lambda}$. We can make this geometry manifest by introducing an *auxiliary* 1-plane $W \in G(1,3)$, and demand that it be orthogonal to $\widetilde{\lambda}$ and that it contains the plane λ. This latter

constraint is equivalent to the somewhat less concise condition that the orthogonal complement of $W^{\perp}$ is orthogonal to λ. Thus, we can represent

$$\mathcal{A}_3^{(1)} = \int \frac{d^{1\times 3}W}{\mathrm{vol}(GL(1))} \frac{\delta^{1\times 4}\big(W\cdot\widetilde{\eta}\big)}{(1)(2)(3)} \delta^{1\times 2}\big(W\cdot\widetilde{\lambda}\big)\delta^{2\times 2}\big(\lambda\cdot W^{\perp}\big), \tag{4.23}$$

where $W\in G(1,3)$ is given by the (1×3)-matrix

$$W \equiv \big(w_1 \;\; w_2 \;\; w_3\big), \tag{4.24}$$

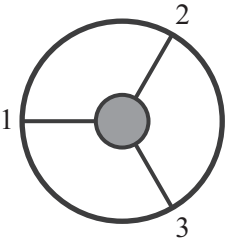
$(a)\equiv \det\{w_a\}$ is a (1×1) 'minor' of the matrix W and $\widetilde{\eta}\equiv(\widetilde{\eta}_1\;\widetilde{\eta}_2\;\widetilde{\eta}_3)$. The δ-function $\delta^{1\times 2}\big(W\cdot\widetilde{\lambda}\big)$ fixes $W\mapsto W^*=\widetilde{\lambda}^{\perp}$ (written above, in (4.19)). On the support of the point $W^*\in G(1,3)$, the remaining δ-functions in (4.23),

$$\delta^{1\times 4}\big(W^*\cdot\widetilde{\eta}\big)\delta^{2\times 2}\big(\lambda\cdot(W^*)^{\perp}\big), \tag{4.25}$$

simply become ordinary super-momentum conservation.

(A small comment is in order here. To make the invariance of the integrand under $GL(1)$ manifest, one has to find a $GL(1)$ invariant way of writing $\delta^{2\times 2}\big(\lambda\cdot W^{\perp}\big)$. As usual, this is achieved by introducing auxiliary variables as explained in detail (and more generality) in section 8.2.)

We can of course make the same generalization for the $\mathcal{A}_3^{(2)}$ vertex:



$$\Leftrightarrow\quad \mathcal{A}_3^{(2)} = \frac{\delta^{2\times 4}(\lambda_1\widetilde{\eta}_1+\lambda_2\widetilde{\eta}_2+\lambda_3\widetilde{\eta}_3)}{\langle 1\,2\rangle\langle 2\,3\rangle\langle 3\,1\rangle}\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big). \tag{4.26}$$

We can think of this as an integral over an auxiliary 2-plane $B\in G(2,3)$ according to

$$\mathcal{A}_3^{(2)} = \int \frac{d^{2\times 3}B}{\mathrm{vol}(GL(2))} \frac{\delta^{2\times 4}\big(B\cdot\widetilde{\eta}\big)}{(12)(23)(31)} \delta^{2\times 2}\big(B\cdot\widetilde{\lambda}\big)\delta^{2\times 1}\big(\lambda\cdot B^{\perp}\big). \tag{4.27}$$

In this case, we can use the constraint $\delta^{2\times 1}\big(\lambda\cdot B^{\perp}\big)$ to localize the integral over B, (somewhat trivially) fixing $B\mapsto B^*=\lambda$, and the minors in the measure trivially become $(1\,2)(2\,3)(3\,1)\mapsto\langle 1\,2\rangle\langle 2\,3\rangle\langle 3\,1\rangle$. As before, the remaining δ-functions in (4.27),

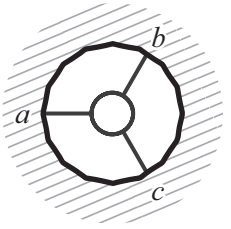
$$\delta^{2\times 4}\big(B^*\cdot\widetilde{\eta}\big)\delta^{2\times 2}\big(B^*\cdot\widetilde{\lambda}\big), \tag{4.28}$$

encode super-momentum conservation.

The crucial feature of these Grassmannian representations of the three-particle amplitudes is that the constraints on the kinematical data λ and $\widetilde{\lambda}$ are now

decoupled and occur *linearly* in the δ-function constraints. This makes it essentially trivial to perform the phase-space integral over the internal lines, making any on-shell graph simply a collection of auxiliary 1-planes $W \in G(1,3)$ and 2-planes $B \in G(2,3)$ associated with the white and black vertices—each carrying with it all the constraints to impose momentum-conservation.

To summarize, for each white vertex involving the (possibly internal) legs (a,b,c) we introduce a 1-plane $W \in G(1,3)$,

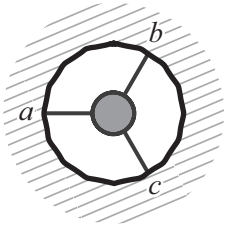


$$\Leftrightarrow \quad W \equiv \left(w_a \; w_b \; w_c\right), \tag{4.29}$$

carrying with it an integration measure,

$$d\Omega_w \equiv \frac{d^{1\times 3}W}{\mathrm{vol}(GL(1))}\frac{1}{(a)(b)(c)}, \tag{4.30}$$

and corresponding constraints; similarly, for each black vertex involving legs (a,b,c) we have a plane $B \in G(2,3)$,



$$\Leftrightarrow \quad B \equiv \left(b_a \; b_b \; b_c\right), \tag{4.31}$$

together with its associated integration measure,

$$d\Omega_b \equiv \frac{d^{2\times 3}B}{\mathrm{vol}(GL(2))}\frac{1}{(a\,b)(b\,c)(c\,a)}, \tag{4.32}$$

and corresponding constraints. Each white vertex imposes *one* relation among $\widetilde{\lambda}$'s:

$$W\cdot\widetilde{\lambda} = w_a\widetilde{\lambda}_a + w_b\widetilde{\lambda}_b + w_c\widetilde{\lambda}_c = 0; \tag{4.33}$$

and each black vertex imposes *two* relations (as the columns b_a of B are 2-vectors):

$$B\cdot\widetilde{\lambda} = b_a\widetilde{\lambda}_a + b_b\widetilde{\lambda}_b + b_c\widetilde{\lambda}_c = 0. \tag{4.34}$$

Thus, for a graph with n_b black vertices, n_w white vertices, and n_I internal edges, we have a total of $2n_b + n_w$ constraints; from these, one constraint is needed to fix (and eliminate) each internal $\widetilde{\lambda}_I$—leaving us with a total of:

$$k \equiv 2n_b + n_w - n_I \tag{4.35}$$

linear constraints relating the external $\widetilde{\lambda}$'s for any given graph. We may write this collection of constraints as $C\cdot\widetilde{\lambda}=0$ for some $(k\times n)$ matrix C, where

$$n = 3n_V - 2n_I, \tag{4.36}$$

with $n_V = n_b + n_w$. Because these are *linear* constraints among the $\widetilde{\lambda}$'s, the matrix C is of course only well defined up to an arbitrary reshuffling of its k equations (a $GL(k)$-transformation of C); and so, C actually represents a point in $G(k,n)$! Of course, integrating out the internal $\widetilde{\eta}$'s follows identically to the $\widetilde{\lambda}$ case, giving us the same final constraints among the external $\widetilde{\eta}$'s as for the $\widetilde{\lambda}$'s.

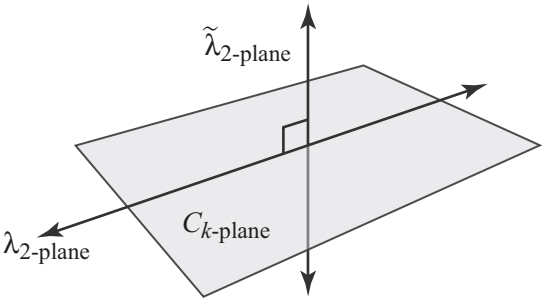
Thus, eliminating the internal $\widetilde{\lambda}_I$ and $\widetilde{\eta}_I$ combines all the "little Grassmannians" $W\in G(1,3)$ and $B\in G(2,3)$ associated with the vertices, and gives us finally a point in the Grassmannian $G(k,n)$ represented by some matrix C that encodes the relations satisfied among the $\widetilde{\lambda}$'s and $\widetilde{\eta}$'s via the δ-functions

$$\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)\delta^{k\times 2}\big(C\cdot\widetilde{\lambda}\big). \tag{4.37}$$

Following the same logic, but exchanging each plane B and W for their orthogonal complements, gives us the complementary set of relations involving the λ's. Not surprisingly, these are simply given by the δ-functions

$$\delta^{2\times(n-k)}\big(\lambda\cdot C^{\perp}\big). \tag{4.38}$$

Geometrically, the ordinary δ-functions constrain the matrix C to be *orthogonal to* $\widetilde{\lambda}$ and to *contain* λ:



(4.39)

It may be worthwhile to clarify the δ-function constraints requiring that C contains λ, $\delta^{2\times(n-k)}\big(\lambda\cdot C^{\perp}\big)$. In particular, in order to see that the complete set of δ-functions is in fact $GL(k)$-invariant, we should clarify how the matrix $C^{\perp}$ is to be precisely defined—it being only defined geometrically up to arbitrary $GL(n-k)$ transformations. Notice that if C *contains* the 2-plane λ, then there must exist a linear combination of the k row vectors of C that *exactly* match λ. In other words, if we parameterize such a linear combination by a $(2\times k)$ matrix ρ, then we should be able to find some ρ for which $\rho\cdot C=\lambda$ (where '$\cdot$' means a sum over the k rows). Therefore, we can rewrite the constraint that C contains λ in terms of this auxiliary

matrix ρ as follows:

$$\delta^{2\times(n-k)}\big(\lambda\cdot C^{\perp}\big)\equiv\int d^{2\times k}\rho\;\delta^{2\times n}\big(\rho\cdot C-\lambda\big). \tag{4.40}$$

Putting everything together, each on-shell diagram is associated with a differential-form obtained by integration over

$$\prod_{\substack{\text{internal}\\ \text{edges } e}}\Big(\frac{1}{\text{vol}(GL(1)_e)}\Big)\prod_w d\Omega_w\prod_b d\Omega_b\;\;\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)\delta^{k\times 2}\big(C\cdot\widetilde{\lambda}\big)\delta^{2\times(n-k)}\big(\lambda\cdot C^{\perp}\big). \tag{4.41}$$

Notice that while freely using the δ-functions to fix each internal λ_I and $\widetilde{\lambda}_I$, we have not modded out by the $GL(1)$-redundancies acting on these momenta (which explains the appearance of the $1/\text{vol}(GL(1))$ factors in (4.41)). It is natural to refer to the net number of auxiliary variables—after modding out by all these $GL(1)$-redundancies—as the *dimension* of the space of configurations $C\in G(k,n)$. As each vertex carries two auxiliary degrees of freedom, and each $GL(1)$ from the internal lines can be used to remove one of them, the 'dimension' associated with an on-shell graph is simply:

$$\dim(C)=2n_V-n_I. \tag{4.42}$$

We should mention that this can be counted in a more direct way from the graph as follows. Because each on-shell graph is trivalent, we have $3n_V=2n_I+n$ so that $\dim(C)=2n_V-n_I=n_I-n_V+n$; and restricting our attention to planar graphs, Euler's formula tells us that $(n_F-n)-n_I+n_V=1$ (where n_F is the number of faces of the graph *including* the n faces of the boundary). Putting these two facts together shows that

$$\dim(C)=n_F-1. \tag{4.43}$$

We will soon see that this is not an accident: there is a natural way in which the degrees of freedom associated with a graph are encoded by its *faces*.

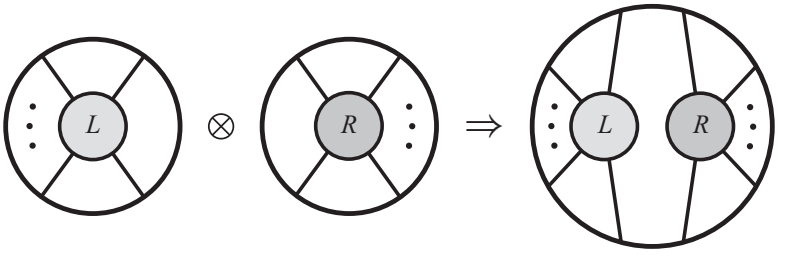
So far, we've described in general terms how to compute the differential-form associated with a given on-shell graph. In the next subsection, we will describe how this can be done systematically using only two very simple, elementary operations; and in section 4.5, we'll show how these two operations can be efficiently automated to construct an explicit representative of the plane C expressed in terms of variables associated with either a graph's *edges* or *faces*.

4.4 Amalgamation of on-shell diagrams

General on-shell diagrams can be built up in steps from more elementary ones using two simple operations: *direct products* and *projections*. Collectively, we

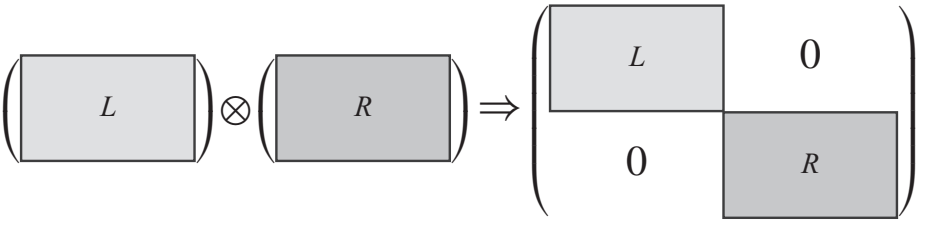
refer to this step-wise construction of more complicated diagrams from simpler ones as *amalgamation* (see [38] for a mathematical construction and [16, 86] for some early steps in the physical setup). In this subsection, we describe both operations in turn, and show how they completely determine the k-plane C associated with any on-shell graph. Since all the $GL(k)$-invariant information about C is given by the ratios of its minors, $(a_1 \cdots a_k)/(b_1 \cdots b_k)$, it suffices for us to simply describe how these two primitive operations act on the minors of C.

The first operation is rather trivial: starting with any two diagrams, we can take their *direct product*:



(4.44)

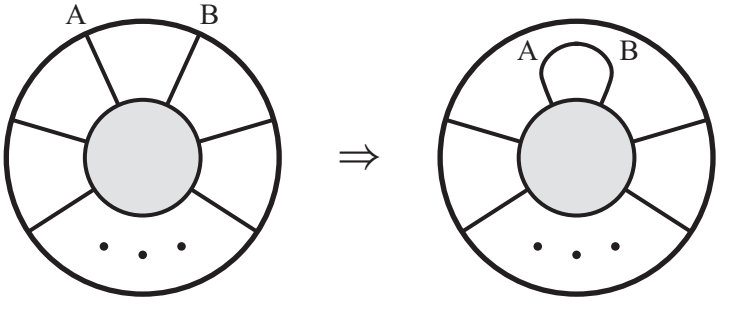
If the left-hand graph is associated with the plane $C_L \in G(k_L, n_L)$, and the right-hand graph is associated with the plane $C_R \in G(k_R, n_R)$, the direct product produces a plane $C_L \otimes C_R \mapsto C \in G(k_L + k_R, n_L + n_R)$ according to



The non-vanishing minors of C are easily expressed in terms of those of C_L and C_R:

$$(a_1 \cdots a_{k_L} b_1 \cdots b_{k_R})\big|_C = (a_1 \cdots a_{k_L})\big|_{C_L} \times (b_1 \cdots b_{k_R})\big|_{C_R}. \tag{4.45}$$

The second operation, *projection*, is more interesting. It corresponds to the identification of two (external) legs—say A and B—of a graph:



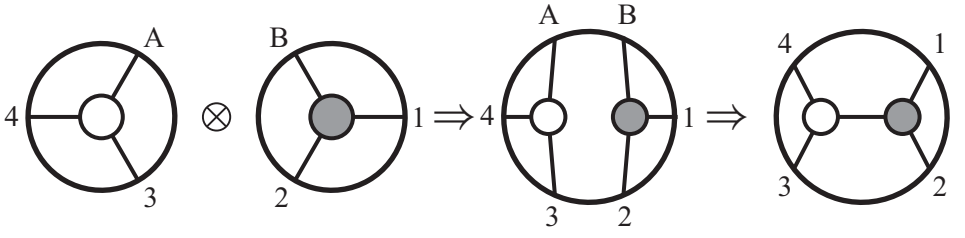
(4.46)

We call this operation "projection" because it takes a plane $C \in G(k+1,n+2)$, and produces a plane $\widehat{C} \in G(k,n)$, which is the *projection of C* onto the quotient of the column space of C modulo $(c_A - c_B)$. This follows directly from how the plane C associated with an on-shell graph is interpreted geometrically as constraints imposed on the external momenta.

For convenience, let us suppose that the $n+2$ particles of the configuration before projection are ordered $(A,B,1,\ldots,n)$. Then the minors of the projection's image $\widehat{C} \in G(k,n)$ will be given in terms of the minors of $C \in G(k+1,n+2)$ according to

$$(a_1 \cdots a_k)|_{\widehat{C}} = (A\, a_1 \cdots a_k)|_C + (B\, a_1 \cdots a_k)|_C \,. \tag{4.47}$$

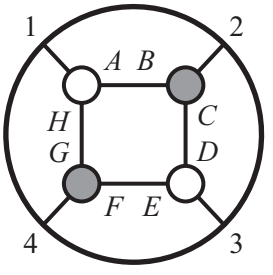
Let us consider a simple case where these two operations are used to construct an on-shell graph. For example, consider the sequence



which builds up the 4-particle factorization graph by first taking the direct product of $W \in G(1,3)$ and $B \in G(2,3)$ to produce a graph associated with a plane $\widehat{C} \in G(3,6)$, then merging legs A and B to produce the final graph associated with a plane $C \in G(2,4)$. As we have described, minors of the final plane $C \in G(2,4)$ are fully specified by those of its constituents; e.g.,

$$\begin{aligned} (13)|_C &= (A13)|_{\widehat{C}} + (B13)|_{\widehat{C}} = 0 + (B1)|_B \times (3)|_W \,; \\ \text{and} \quad (24)|_C &= (A24)|_{\widehat{C}} + (B24)|_{\widehat{C}} = 0 + (B2)|_B \times (4)|_W \,. \end{aligned} \tag{4.48}$$

Let us look at one more interesting example: the amalgamation of diagrams generating the 4-particle tree-amplitude:

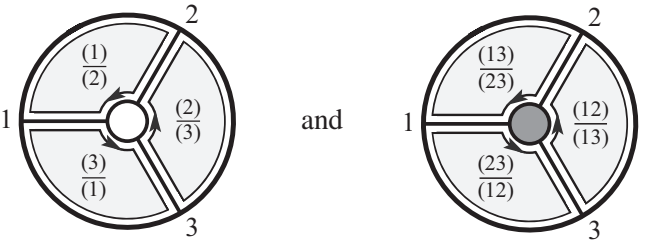


(4.49)

Following the amalgamation rules described above, we find, for example that

$$\frac{(24)}{(13)} = \left(\frac{(F4)(H)}{(GF)(1)}\right)\left(\frac{(B2)(D)}{(CB)(3)}\right) + \left(\frac{(C2)(A)}{(BC)(1)}\right)\left(\frac{(G4)(E)}{(GF)(3)}\right). \tag{4.50}$$

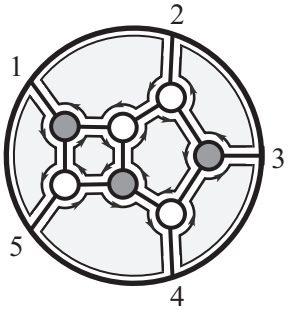
Notice that the amalgamation picture makes it clear that C will only depend on special combinations of the minors of the matrices associated with its constituent vertices. This ultimately stems from the fact that the only $GL(k)$-invariant data associated with the vertices themselves are the ratios of minors. These appear, for example, as the face variables of the three-particle diagrams:



(4.51)

Here, we have used arrows to show how the ratios transform under the little group.

Now, a very simple but important observation is that the final point in $G(k,n)$ obtained from amalgamation must obviously be completely invariant under the little-group rescaling of any internal line. This means that only combinations of minors that are invariant under these scaling are ultimately relevant to our description of C. Graphically, it is clear that these are given by products of such ratios, as following along the boundary of a face we form a closed path. A face variable, then, can be built as the product of these variables along its boundary. To illustrate the point, consider the following graph—associated with a generic plane in $G(2,5)$:



(4.52)

Thus, while the variables describing the matrix C can be constructed from the variables of the planes B and W attached to each vertex, we may alternatively view C as being described by variables f_i associated with its *faces*. Note that the product of the face variables for each $G(1,3)$–$G(2,3)$ vertex is manifestly equal to 1 (see (4.51)); and so, it is easy to see that there are only *two* independent degrees of freedom per vertex—matching our calculation that $\dim(C) = n_F - 1$. This clearly persists to larger diagrams, ensuring that $\prod_i f_i = 1$, which always accounts for the "minus 1" in the formula for the dimension of C. And so, the degrees of freedom are all but one of the face variables, say f_*. Rescaling $f_i \mapsto \widehat{f_i} \equiv f_i/f_*$, the integration measure (4.41) for the auxiliary parameters in C becomes simply

$$\prod_{\substack{\text{internal}\\ \text{edges } e}} \Big(\frac{1}{\text{vol}(GL(1)_e)}\Big) \prod_w d\Omega_w \prod_b d\Omega_b = \pm \prod_{\substack{\text{rescaled}\\ \text{faces } \widehat{f}_i}} \frac{d\widehat{f}_i}{\widehat{f}_i}. \tag{4.53}$$

Notice the '±' in (4.53). Because the volume is *oriented*—e.g., it is anti-symmetric with respect to reordering the canonical coordinates—volume forms constructed out of different (ordered) sets of canonical coordinates can be oppositely oriented. We will mostly leave this issue aside until Chapter 12, where we will describe a general procedure for determining the relative orientations of volume forms given any pair of canonical coordinate charts.

But before moving on, we should mention how the choice of f_* affects the orientation of the resulting coordinates. If we denote the face variables $\vec{f} \equiv \{f_0, \ldots, f_d\}$, then the volume form may be written in terms of these coordinates as

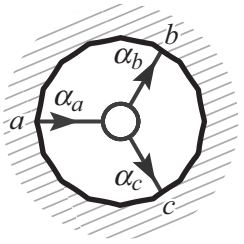
$$\begin{aligned} \Omega_\sigma\big(\vec{f}\big) &\equiv \Big(\frac{df_0}{f_0} \wedge \cdots \wedge \frac{df_d}{f_d}\Big)/\text{vol}(GL(1)), \\ &\equiv (-1)^i \Big(\frac{df_0}{f_0} \wedge \cdots \wedge \frac{df_{i-1}}{f_{i-1}} \wedge \frac{df_{i+1}}{f_{i+1}} \wedge \cdots \wedge \frac{df_d}{f_d}\Big); \end{aligned} \tag{4.54}$$

which makes explicit how the choice of f_* affects the volume form's orientation. A much more detailed discussion on the (relative) orientations of positroid volume forms for different sets of canonical coordinates can be found in Chapter 12.

4.5 "Boundary measurements" and canonical coordinates

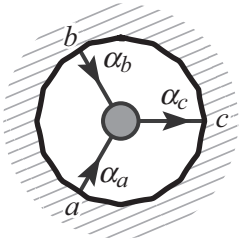
Let us now turn to the problem of explicitly determining a matrix representative C associated with a given on-shell graph. We will first do this in a very efficient—but somewhat overly redundant—way by attaching variables α_e to all the *edges* of a graph; and then, we will see how this procedure can be translated (with less redundancy) in terms of variables attached to a graph's *faces*.

One strategy for explicitly constructing the k-plane C encoding the system of constraints (4.39) associated with an on-shell graph is to set the degrees of freedom associated with each vertex in a way that allows us to eliminate all internal momenta as efficiently as possible. Of course, each vertex carries with it only two degrees of freedom. But it turns out to be useful to introduce an additional $GL(1)$-redundancy at each vertex, so that every *leg* attached to a given vertex carries its own degree of freedom (making it easier to pair up the degrees of freedom attached to internal lines between vertices). To further simplify the elimination of internal momenta from the ultimate system of equations relating the $\widetilde{\lambda}$, it will be helpful also to provide an orientation to each edge, so that each white (black) vertex has one (two) edges directed inward. With these decorations, each white vertex corresponds to:



$$\Leftrightarrow W \equiv \overbrace{\begin{pmatrix} a & b & c \\ \alpha_a^{-1} & -\alpha_b & -\alpha_c \end{pmatrix}} \implies \widetilde{\lambda}_a = \alpha_a(\alpha_b \widetilde{\lambda}_b + \alpha_c \widetilde{\lambda}_c); \tag{4.55}$$

and each black vertex corresponds to:



$$\Leftrightarrow B \equiv \overbrace{\begin{pmatrix} a & b & c \\ \alpha_a^{-1} & 0 & -\alpha_c \\ 0 & \alpha_b^{-1} & -\alpha_c \end{pmatrix}} \Rightarrow \left\{ \begin{array}{l} \widetilde{\lambda}_a = \alpha_a \alpha_c \widetilde{\lambda}_c \\ \widetilde{\lambda}_b = \alpha_b \alpha_c \widetilde{\lambda}_c \end{array} \right\}. \tag{4.56}$$

Decorating a graph in this way is called giving it a *perfect orientation*; and it is a general fact that all two-colored, trivalent graphs *relevant to physics* with supersymmetry can be given a perfect orientation.

(The only graphs that cannot be given a perfect orientation are those containing a subgraph with $k \leq 0$ or $k \geq \nu$ (where ν denotes the number of legs of the *sub*graph). This obstruction is closely tied to an inability to eliminate some internal line's λ_I or $\widetilde{\lambda}_I$ from the complete system of equations. But this subtlety plays no role in our story, as the differential form associated with such a graph *always* vanishes due to the $\widetilde{\eta}_I$ integration. And so, these 'pathological' diagrams never contribute to physically relevant processes.)

Once we have given a perfect orientation, the system of equations $C \cdot \widetilde{\lambda}$ becomes trivial to construct: each vertex can be viewed as giving an equation that expands the $\widetilde{\lambda}$'s of the vertex's *sources* in terms of those of its *sinks*. Combining all such equations then gives us an expansion of the external sources' $\widetilde{\lambda}$'s in terms of those of the external sinks. Notice that when identifying two legs (I_{in}, I_{out}) during amalgamation, the degree of freedom lost in the process is accounted for via the replacement of the pair $(\alpha_{I_{in}}, \alpha_{I_{out}})$ with the single variable $\alpha_I \equiv \alpha_{I_{in}} \alpha_{I_{out}}$.

If we denote the external sources of a graph by $\{a_1, \ldots, a_k\} \equiv A$, then the final linear relations imposed on the $\widetilde{\lambda}$'s can easily be seen to be given by

$$\widetilde{\lambda}_A + c_a^A \widetilde{\lambda}_a = 0, \tag{4.57}$$

with

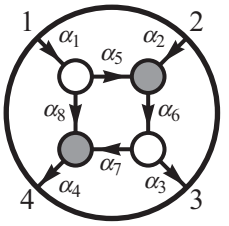
$$c_a^A = -\sum_{\Gamma \in \{A \rightsquigarrow a\}} \prod_{e \in \Gamma} \alpha_e \,, \tag{4.58}$$

and where $\Gamma \in \{A \rightsquigarrow a\}$ is any (directed) path from A to a in the graph. (If there is a closed, directed loop, then the geometric series should be summed—we will see

an example of this in (4.65).) The entries of the matrix c_a^A are called the "boundary measurements" of the on-shell graph. The on-shell form on $C(\alpha) \in G(k,n)$ can then be written in terms of the variables c_a^A according to:

$$\left(\prod_{\text{vertices } v} \frac{1}{\text{vol}(GL(1)_v)}\right)\left(\prod_{\text{edges } e} \frac{d\alpha_e}{\alpha_e}\right)\delta^{k\times 4}(C\cdot\widetilde{\eta})\delta^{k\times 2}(C\cdot\widetilde{\lambda})\delta^{2\times(n-k)}(\lambda\cdot C^{\perp}). \quad (4.59)$$

Let us consider a simple example to see how this works. Consider the following perfectly oriented graph:

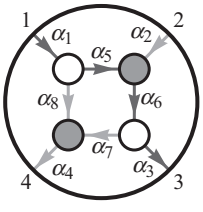


(4.60)

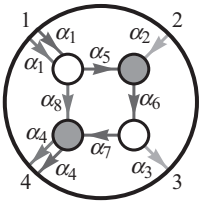
Using the equations for each directed 3-particle vertex, we can easily expand the $\widetilde{\lambda}$ of each source—legs 1 and 2—in terms of those of the sinks—legs 3 and 4; e.g.,

$$\widetilde{\lambda}_2 = \alpha_2\alpha_6(\alpha_3\widetilde{\lambda}_3 + \alpha_7(\alpha_4\widetilde{\lambda}_4)). \quad (4.61)$$

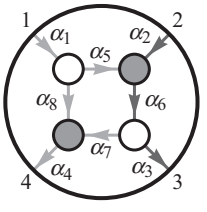
Such expansions obviously result in (4.58): the coefficient c_a^A of $\widetilde{\lambda}_a$ in the expansion of $\widetilde{\lambda}_A$ is simply (minus) the product of all edge-variables α_e along any path $\Gamma \in \{A \rightsquigarrow a\}$. Doing this for all the c_a^A of our example above, we find that



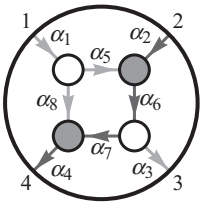
$-c_3^1 = \alpha_1\alpha_5\alpha_6\alpha_3$



$-c_4^1 = \alpha_1\alpha_5\alpha_6\alpha_7\alpha_4 + \alpha_1\alpha_8\alpha_4$



$-c_3^2 = \alpha_2\alpha_6\alpha_3$



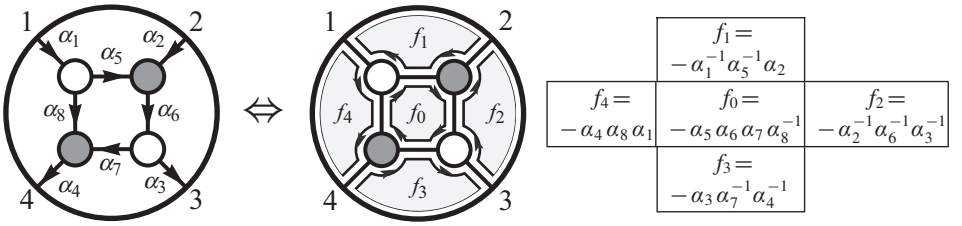
$-c_4^2 = \alpha_2\alpha_6\alpha_7\alpha_4$

Thus, the final relations involving the $\widetilde{\lambda}$'s is encoded by the matrix $C \equiv \begin{pmatrix} 1\,0\,c_3^1\,c_4^1 \\ 0\,1\,c_3^2\,c_4^2 \end{pmatrix}$.

Notice that only certain combinations of edge-weights appear in the equations. This happens for a very simple—and by now familiar—reason. Think of the $GL(1)$-redundancy of each vertex as a gauge group, with the variable of a directed edge charged as a "bi-fundamental" of the $GL(1) \times GL(1)$ of the vertices it connects. Since the configuration C must be invariant under these "gauge groups," only gauge-invariant combinations of the edge variables can appear. And just as we saw in the previous subsection, these combinations are those familiar from lattice gauge theory and can be viewed as encoding the flux though each closed loop in the graph—that is, each of its faces. Fixing the orientation of each face

to be clockwise, the flux through it is given by the product of α_e (α_e^{-1}) for each aligned (anti-aligned) edge along its boundary. For future convenience, we define the face *variables* f_i to be *minus* this product.

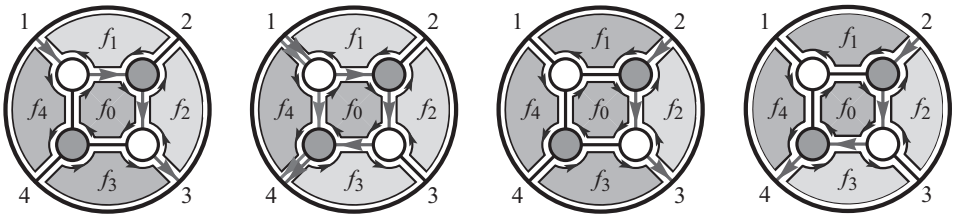
Applying this to the example above, we find:



The boundary measurements c_a^A can then be expressed in terms of the faces by

$$c_a^A = -\sum_{\Gamma\in\{A\rightsquigarrow a\}}\prod_{f\in\widehat{\Gamma}}(-f)\,, \tag{4.62}$$

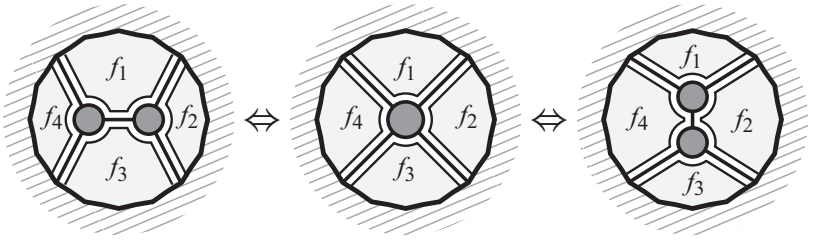
where $\widehat{\Gamma}$ is the *clockwise* closure of Γ. (If there are any closed, directed loops, the geometric series of faces enclosed should be summed.) The faces of course overcount the degrees of freedom by one, and this is reflected by the fact that $\prod_i(-f_i) = 1$.



4.6 Coordinate transformations induced by moves and reduction

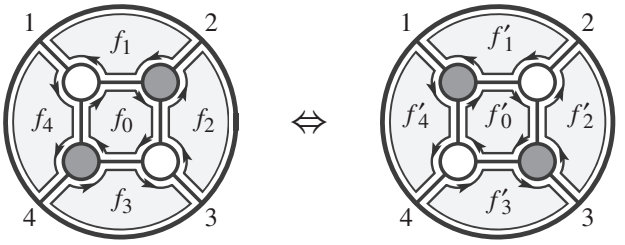
Let us now examine how the identification of diagrams via merge-operations, square moves, and bubble-deletion is reflected in the coordinates—the edge- or face-variables—used to parameterize cells $C \in G(k,n)$. As usual, the simplest of these is the merge/unmerge operation, which trivially leaves any set of coordinates unchanged. For example, in terms of the face variables, it is easy to see

that



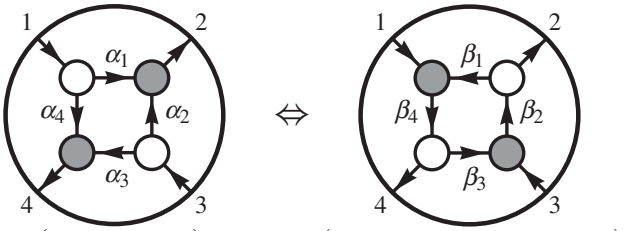
(4.63)

The square move is more interesting. It is obvious that squares with opposite coloring both give us a generic configuration in $G(2,4)$, but (as we will soon see) the square move acts rather nontrivially on coordinates used to parameterize a cell,



(4.64)

Let us start by determining the precise way the face-variables f_i and f'_i of square move related diagrams are related to one another. To do this, we will provide perfect orientations (decorated with edge variables) for both graphs, allowing us to compare the resulting boundary-measurement matrices in each case. Because these two boundary-measurement matrices must represent the same point in $G(2,4)$, we will be able to determine explicitly how all the various coordinate charts are related—including the relationship between the variables f_i and f'_i. Our work will be considerably simplified if we remove the $GL(1)$-redundancies from each vertex, leaving us with a non-redundant set of edge-variables. Of course, *any* choice of perfect orientations for the graphs, and any fixing of the $GL(1)$-redundancies, would suffice for our purposes; but for the sake of concreteness, let us consider the following:



(4.65)

$$\begin{pmatrix} 1 & -\alpha_1 & 0 & -\alpha_4 \\ 0 & -\alpha_2 & 1 & -\alpha_3 \end{pmatrix} \qquad \begin{pmatrix} 1 & -\beta_2\beta_3\beta_4\Delta & 0 & -\beta_4\Delta \\ 0 & -\beta_2\Delta & 1 & -\beta_1\beta_2\beta_4\Delta \end{pmatrix}$$

Here, we have written the matrices $C(\alpha)$ and $C(\beta)$ obtained as boundary measurements as discussed in section 4.5. The factor Δ in $C(\beta)$ is given by

$$\Delta \equiv \frac{1}{1-\beta_1\beta_2\beta_3\beta_4}, \tag{4.66}$$

and arises from summing the infinite geometric series of paths that circle around the internal loop of the perfectly oriented graph. The edge-variables in (4.65) used as coordinates in $G(2,4)$ are closely-related to the face-variables in (4.64).

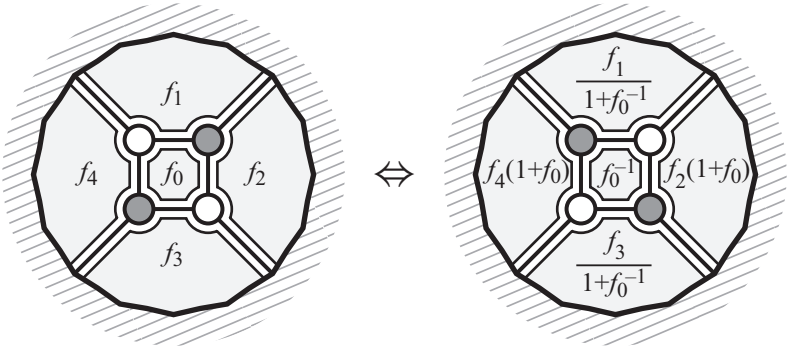
It is not hard to express the face variables in terms of the edge variables for the two orientations in (4.64). It is easy to see that

$$\begin{array}{l} f_0 = -\alpha_1\alpha_2^{-1}\alpha_3\alpha_4^{-1},\ f_1 = -\alpha_1^{-1},\ f_2 = -\alpha_2,\ f_3 = -\alpha_3^{-1},\ f_4 = -\alpha_4^{-1}; \\ f_0' = -(\beta_1\beta_2\beta_3\beta_4)^{-1},\ f_1' = -\beta_1\ ,\ f_2' = -\beta_2,\ f_3' = -\beta_3\ ,\ f_4' = -\beta_4\ . \end{array} \tag{4.67}$$

Because the boundary measurements must represent the same point in the Grassmannian regardless of whether we use α or β coordinates, we see that

$$\left\{\begin{array}{l} \alpha_1 = \beta_2\beta_3\beta_4\Delta \\ \alpha_2 = \beta_2\Delta \\ \alpha_3 = \beta_1\beta_2\beta_4\Delta \\ \alpha_4 = \beta_4\Delta \end{array}\right\} \Rightarrow \left\{\begin{array}{ll} -\beta_1 = f_1' = -\alpha_2^{-1}\alpha_3\alpha_4^{-1}\Delta & = f_1 f_0 \Delta \\ -\beta_2 = f_2' = -\alpha_2\Delta^{-1} & = f_2\Delta^{-1} \\ -\beta_3 = f_3' = -\alpha_1\alpha_2^{-1}\alpha_4^{-1}\Delta & = f_3 f_0 \Delta \\ -\beta_4 = f_4' = -\alpha_4\Delta^{-1} & = f_4\Delta^{-1} \\ \quad \therefore\ \ f_0' = -\alpha_1^{-1}\alpha_2\alpha_3^{-1}\alpha_4 = f_0^{-1} & \end{array}\right\}. \tag{4.68}$$

Observing that $\Delta = (1+f_0'^{-1})^{-1} = (1+f_0)^{-1}$, we therefore conclude that a square move alters face-variables according to the following:



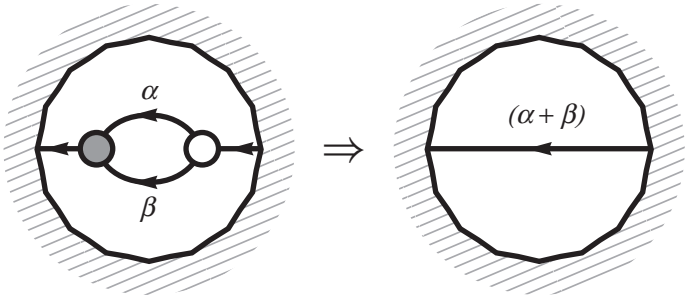
(4.69)

This transformation of the face variables is an example of a more general operation related to *cluster transformations* as described in section 16.2. Note that, crucially, our form is invariant under this transformation (up to an overall change of sign):

$$\prod_f \frac{df}{f} = -\prod_{f'} \frac{df'}{f'}. \tag{4.70}$$

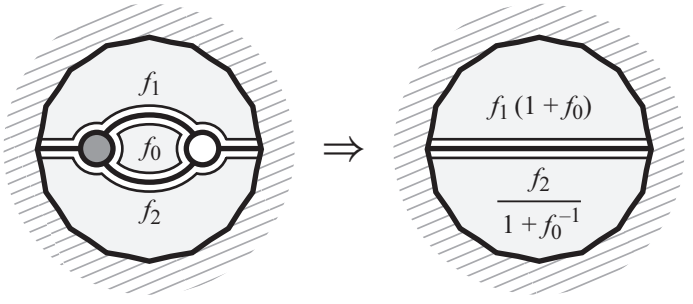
The invariance of the measure (modulo an overall sign) guarantees that the on-shell forms associated with diagrams related by square moves are the same—differing only by a change of coordinates used.

Let us now turn to bubble-deletion. It is easy to see that the following oriented subdiagrams always lead to exactly the same boundary measurements:



(4.71)

Following the same logic used to analyze the square move, we find that the face-variables of these two diagrams are related as follows:



(4.72)

Note again the crucial fact that the measure is invariant under this transformation (again, up to an overall sign):

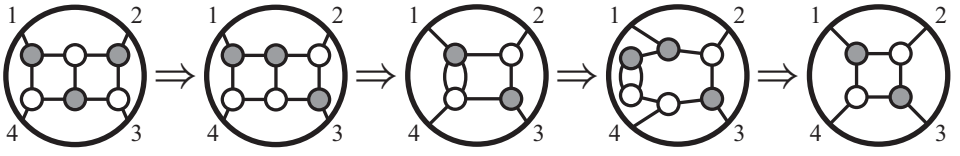
$$\frac{df_0}{f_0} \wedge \frac{df_1}{f_1} \wedge \frac{df_2}{f_2} = -\frac{df_0'}{f_0'} \wedge \frac{df_1'}{f_1'} \wedge \frac{df_2'}{f_2'}, \tag{4.73}$$

where $f_0' = f_0^{-1}$. The change of variables from $f \to f'$ eliminates all dependence on f_0 associated with the bubble from the final point in the Grassmannian. Of course, the variable f_0 *remains in the measure*, but it cleanly factors out as an overall prefactor of $d\log(f_0)$. As we will see later on, MHV amplitude integrands—to all loop-orders—are always the tree-amplitude, dressed with many additional $d\log$-factors arising from bubble-deletion. These "irrelevant" factors in the measure encode the internal degrees of freedom of the loop-momenta.

If instead of the *integrand* for scattering amplitudes, we were interested in the *residues* of the on-shell differential form—to compute, e.g. "leading singularities"—then these "irrelevant" $d\log$-factors *really are irrelevant*: *any* residue involving them will give either one or zero.

Due to reduction, then, the number of interesting *residues* of general (non-reduced) on-shell diagrams turns is in fact finite despite the seemingly infinite number of possible diagrams. Notice that in our way of thinking about 'leading singularities' and on-shell diagrams, we've made no distinction whatsoever between what have historically been called "ordinary" versus "composite" objects [87, 88]. Historically, *reducible* on-shell diagrams were those with "irrelevant" additional degrees of freedom that could be systematically trivialized away.

One example of such an on-shell form is the 'double box' involving four particles; this on-shell diagram has been known to include *one* unfixed degree of freedom that factorizes out of the diagram trivially upon bubble-deletion:



(4.74)

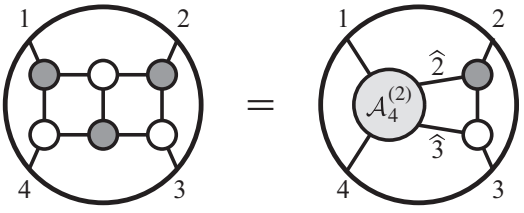
As discussed in generality above, the variable "lost" during bubble-deletion is in reality just a bare $d\log(\alpha)$ in the measure.

4.7 Relation to composite leading singularities

When all the auxiliary degrees of freedom of an on-shell form can be localized by kinematical constraints, we can think of it as having been obtained by starting with the (n_F-n)-loop integrand for the scattering amplitude, and successively putting (off-shell) Feynman propagators on-shell ('cutting' them) until the on-shell diagram is obtained. Such on-shell diagrams are referred to as "leading singularities." Thought of in this way, they are secondary—derived—quantities obtained from the 'primary' object, the loop integrand. An important physical point of our present work (discussed more thoroughly in Chapter 17) is that it is much more fruitful to take the opposite viewpoint: that 'loop integrands' are in fact 'derived' from on-shell diagrams. However, since the concept of a "leading singularity" will likely be more familiar to most readers, in this section we will briefly review how leading singularities have been used to inform us about scattering amplitudes, and discuss in particular the subtle issue of *composite* leading singularities—which is closely related to reducibility. (This discussion is meant only to make contact with this point in previous literature, and isn't especially germane to the rest of our book.)

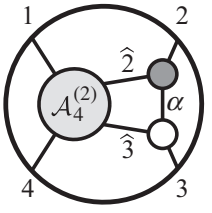
The reduction procedure is related to what was called the "computation of composite leading singularities" in the physics literature [87–90] (see [91–93] for

recent developments). In order to make the connection between the modern and the old procedures transparent, let us explain what a composite leading singularity means for the four-point example already examined above. Starting with the diagram with two faces, one realizes that any of the two squares actually represents a full four-particle tree amplitude. Choose the left one for example and draw the equivalent figure,



(4.75)

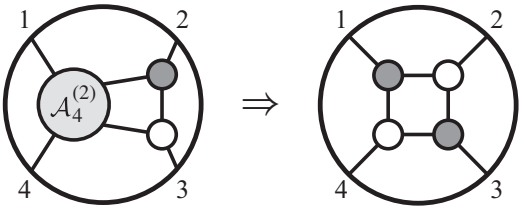
At this point the attentive reader can recognize this as a BCFW-bridge on a physical scattering amplitude; it is given by the differential form



$$= \frac{d\alpha}{\alpha}\mathcal{A}_4^{(2)}(\alpha), \tag{4.76}$$

where the α-dependence of $\mathcal{A}_4^{(2)}$ results from that of the shifted momenta $\widehat{2}$ and $\widehat{3}$. This on-shell form has only two poles in α: a trivial pole at $\alpha = 0$, and another where the $\mathcal{A}_4^{(2)}$ factorizes. Of course, as there are *only two* poles in the α-plane, their residues sum to zero, and hence differ only by a sign; as the $\alpha = 0$ residue is manifestly the *undeformed* tree amplitude $\mathcal{A}_4^{(2)}(\alpha = 0)$, so is the other (up to a sign).

The composite leading singularity technique was based on the observation that the pole at $(p_1 + p_{\widehat{2}})^2 = 0$ is guaranteed to be there simply as a pole of the physical $\mathcal{A}_4^{(2)}(\alpha)$ tree amplitude. Therefore the pole at $(p_1 + p_{\widehat{2}})^2 = 0$, in combination with the other three on-shell conditions on the loop momenta already in the figure, can be used to determine a residue. This gives rise to



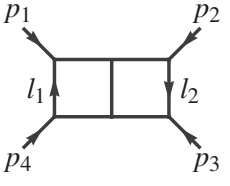
(4.77)

which is none other than the on-shell diagram for a four-point amplitude $\mathcal{A}_4^{(2)}$.

We note in passing that this gives yet another ideal use of bubbles. Suppose that one is given an on-shell diagram corresponding to a leading singularity—namely, an on-shell diagram that evaluates to an algebraic function of external momenta (conditions for this to happen are discussed in Chapter 11). Next, apply a BCFW-bridge to the diagram and ask what its possible poles and corresponding residues are as a function of the BCFW variable α. Let us again return to discussing the same four-particle example. We can ask how we could have known that there was a pole in the '$s_{12}(\alpha) \to 0$' channel and not in any other channel, by only manipulating the graph. The answer is already apparent in (4.74): find a bubble and the channel of the bubble becomes the pole required by unitarity!

Composite leading singularities were first developed in order to compute two-loop amplitudes following a technique that was very successful at one loop [94]. While Feynman diagrams are even hard to write down explicitly for loop amplitudes, it is known that loop integrals can be reduced to a linear combination of basic standard integrals [95]. The idea is then to start with the most general linear combination of such basic integrals and find ways of computing the coefficients. This is known as the "unitarity-based method" [96–100] (for recent applications of these techniques, see e.g. [93, 101]). In more modern language, the key idea is to use contour integrals to compute the coefficients. At one loop, $\mathcal{N} = 4$ super Yang–Mills only requires integrals with four propagators. Thus, the four-dimensional contour for computing a given coefficient is then obviously defined by the four propagators of the given integral.

At two loops and four particles the basis of integrals must include one such as:



(4.78)

Now there are eight integration variables but only seven propagators. Naïvely it seems that this integral does not have any non-vanishing residues with codimension 8. The key observation is that the propagators are nonlinear functions of the integration variables and therefore computing the ℓ_1 integral using the T^4 contour defined by the left box gives rise to $1/\big(s_{12}(\ell_2)s_{41}\big)$, which is ℓ_2-dependent. This can then be used together with the three propagators already present on the right to define a second T^4 contour and hence a non-vanishing residue. The ℓ_2-dependent pole, $1/\big(s_{12}(\ell_2)\big)$, generated in this form is precisely what is needed for the new computation to be that of a single scalar box on-shell diagram.

In this way of thinking about things, the existence of composite residues is unexpected, and is made possible by "hidden" poles that are produced by Jacobian factors that appear as residues are taken. In our new picture, *all* the singularities are manifestly exposed in our "*d*log" measure for edge or face variables. There is no distinction between "composite" and "ordinary" singularities, and they are all treated together in a systematic and unified way.

5

Configurations of vectors and the *positive* Grassmannian

We have seen that every on-shell graph is associated with a $(k\times n)$ matrix C, where a reduced graph with n_F faces gives us an (n_F-1)-dimensional submanifold of the Grassmannian $G(k,n)$. We have also seen that the invariant content of an on-shell diagram is given by the permutation that labels it. We will now link these two observations by showing that the submanifold in the Grassmannian associated with an on-shell graph is also characterized—for geometric reasons—by the *same* permutation which labels the graph.

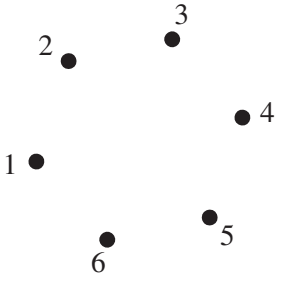
Our discussion will be most transparent if we think of the Grassmannian in a complementary way to our presentation so far: instead of viewing the $(k\times n)$ matrix C *horizontally*, as a k-plane spanned by its rows, we want to now view C *vertically*—as a collection of n, k-dimensional columns. The $GL(k)$-invariant data to describe any configuration are ratios of minors: $(a_1\cdots a_k)/(b_1\cdots b_k)$. Intuitively, a *generic* plane C would be one for which none of its minors vanish. Such a configuration would have $k(n-k)$ degrees of freedom. The vanishing of any minor of C implies some linear dependence among its columns. Allowing for all possible linear dependencies among the columns of C leads to the "matroid stratification" [102] of configurations, which is known to be arbitrarily complicated. (Indeed, it was proven in [103] that *all* algebraic varieties are part of this matroid stratification, so understanding this amounts to *completely* taming the entire category of algebraic varieties!) However, if we impose one small restriction on the set of admissible linear dependencies, we will find that a rich, simple, and very beautiful structure emerges.

5.1 The geometry and combinatorics of the positroid stratification

Notice that any configuration C associated with an on-shell, planar graph is endowed with a cyclic ordering for the columns $\{c_1,\ldots,c_n\}$. It is therefore natural to consider a stratification of $G(k,n)$ that involves only linear dependencies among

(cyclically) *consecutive* chains of columns. This is known as the *positroid stratification* [39, 40] (see also [34, 104])—the cells of which are called *positroids*—and will turn out to be precisely what is relevant to the physics of on-shell diagrams. In order to understand the connection most clearly, we will first discuss the stratification in some detail on its own, and show how these configurations are characterized by permutations. We will then see how the geometrically defined permutation that characterizes C is precisely the one that would label the graph.

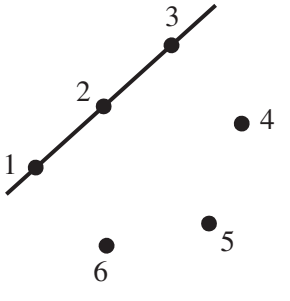
Before describing the stratification generally, it may help to consider some simple examples. Since the kinematical data describing the external particles enjoy a rescaling symmetry, we often find it useful to transfer this symmetry to the columns of C, identifying $c_a \sim t_a c_a$, so that (non-vanishing) columns c_a can be thought of as elements in $\mathbb{P}^{(k-1)}$ (vanishing columns simply being absent from the space). This makes it a little easier to visualize configurations—at least for small k. Consider a generic configuration $C \in G(3,6)$, whose six columns—viewed as points in $\mathbb{P}^2$—are arranged according to:



(5.1)

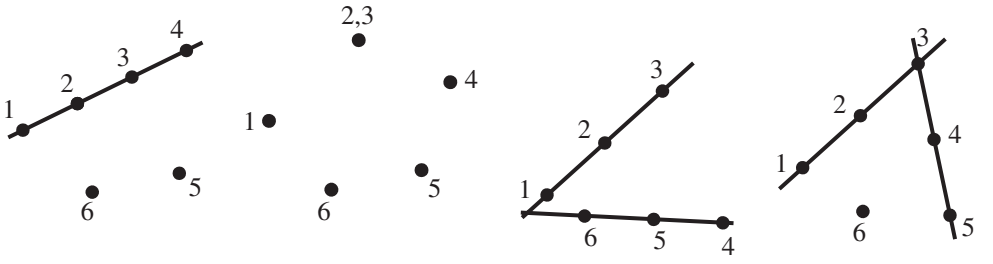
As no three of the columns are linearly dependent, this indeed represents a generic configuration in $G(3,6)$, and has $3(6-3)=9$ degrees of freedom.

The simplest *consecutive* constraint we could impose on (5.1) would be to force any three consecutive columns to become linearly dependent—projectively, collinear. For example, we could require that the minor (123) vanish:

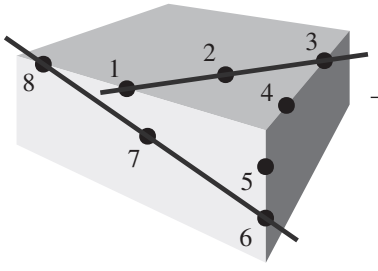


(5.2)

From this configuration, seven possible further restrictions are possible, including:



For $k \leq 3$, it is easy to describe such configurations geometrically—being easily visualizable. But such geometric descriptions rapidly become cumbersome as k increases: even for $k=4$—which is still possible to visualize in three-dimensional space—configurations obtainable using only consecutive constraints can become impressively complex. Consider for example the following configuration in $G(4,8)$:



consec. chains of columns	span
(1) (2) (3) (4) (5) (6) (7) (8)	$\mathbb{P}^0$
(123) (34) (45) (56) (678) (81)	$\mathbb{P}^1$
(56781) (81234) (3456)	$\mathbb{P}^2$

(5.3)

A more systematic way to describe any configuration in this stratification would be to list the *ranks* of spaces spanned by all consecutive chains of columns. Labeling columns mod n, let us define

$$r[a;b] \equiv \text{rank}\{c_a, c_{a+1}, \ldots, c_b\}; \tag{5.4}$$

then knowing $r[a;b]$ for all n^2 pairs of columns $a \leq b$ would suffice to reconstruct any particular configuration. This data is obviously highly redundant: for example, $r[a;a+n-1]=k$ for all a. We can discover how this data can be encoded more efficiently if by first organizing it in a clever way (we thank Pierre Deligne for suggesting this to us):

$$
\begin{array}{|ccccccccccc|}
\hline
 & & & & & & & & r[\ n\ ;2n-1] & \cdots & 2n-1 \\
 & & & & & & r[n-1;2n-2] & & \vdots & \ddots & 2n-2 \\
 & & & & & \ddots & \vdots & & \vdots & \ddots & \vdots \\
 & & & r[\ 2\ ;n+1] & \cdots & r[n-1;n+1] & & r[\ n\ ;n+1] & & \ddots & n+1 \\
r[\ 1\ ;\ n\ ] & & & \vdots & \cdots & r[n-1;\ \ n\ ] & & r[\ n\ ;\ \ n\ ] & & & n \\
\vdots & & & \vdots & \cdots & r[n-1;n-1] & & & & & n-1 \\
r[\ 1\ ;\ 3\ ] & r[\ 2\ ;\ 3\ ] & \ddots & & & & & & & & \vdots \\
r[\ 1\ ;\ 2\ ] & r[\ 2\ ;\ 2\ ] & & & & & & & & & 2 \\
r[\ 1\ ;\ 1\ ] & & & & & & & & & & 1 \\
\hline
1 & & 2 & \cdots & & n-1 & & & n & \cdots &
\end{array}
\tag{5.5}
$$

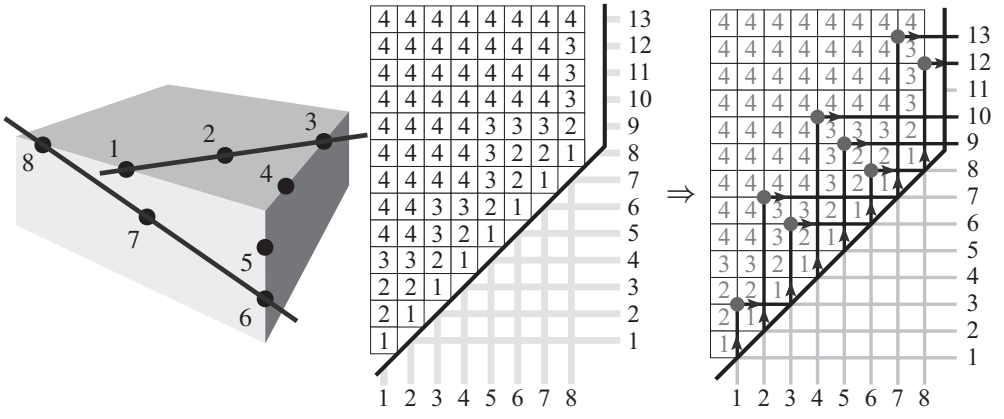
The advantages of arranging the ranks in this way will become clear momentarily. Notice that for each pair of adjacent columns in the table above $(a\,a+1)$ there is some b sufficiently large such that $r[a;b]=r[a+1;b]$, as $r[a;b]$ is bounded above by k and strictly increases with b (moving vertically in (5.5)). Moreover, it is easy to see that if $r[a;b]=r[a+1;b]$ for some b, then $r[a;c]=r[a+1;c]$ for all $c\geq b$, as we would have $c_a\in\text{span}\{c_{a+1},\ldots,c_b\}$, and so $\text{span}\{c_a,\ldots,c_b\}\subset\text{span}\{c_a,\ldots,c_c\}$ for all $c\geq b$. The same argument shows that, moving from right to left along each pair of consecutive rows in (5.5), for any c there exists a b such that $r[b;c]=r[b;c+1]$, and that for all $a<b$, $r[a;c]=r[a;c+1]$.

Because $r[a;b]\geq r[a+1;b]$ in general, for each a there must be a *nearest* column, which we will denote (suggestively) by '$\sigma(a)$'—with $\sigma(a)\geq a$—such that $r[a;\sigma(a)]=r[a+1;\sigma(a)]$. Notice that this implies that $r[a;\sigma(a)]=r[a;\sigma(a)-1]>r[a+1;\sigma(a)-1]$, as otherwise $\sigma(a)$ would not be the nearest. Similarly, we see that a must be the *maximal* column $a\leq\sigma(a)$ such that $r[a;\sigma(a)]=r[a;\sigma(a)-1]$. Thus, there is a *unique* point vertically along each pair of consecutive columns and a *unique* point horizontally along each pair of consecutive rows, where the table locally looks like this:

$$
\begin{array}{|c|c|}
\hline
r[\ a\ ;\ \sigma(a)\ ] & r[a+1\,;\ \sigma(a)\ ] \\
\hline
r[\ a\ ;\sigma(a)-1] & r[a+1\,;\sigma(a)-1] \\
\hline
\end{array}
\quad\Leftrightarrow\quad
\begin{array}{|c|c|}
\hline
r & r \\
\hline
r & r-1 \\
\hline
\end{array}
\ \sigma(a)
\qquad
\begin{array}{c}
\\ \\ \\ a
\end{array}
\tag{5.6}
$$

These "hooks" show that σ is in fact a *permutation* among the labels $\{1,\ldots,n\}$ of the column vectors. Actually, because this definition of σ differentiates between $\sigma(a)=a$ (which occurs whenever $r[a;a]=0$) and $\sigma(a)=a+n$, σ is automatically a *decorated* permutation as defined in section 3.1.

We can see how the permutation encoded by these hooks can be read off from the table of ranks, (5.5), by considering the example configuration given above, (5.3):



(This picture of the permutation σ is similar to the "juggling patterns" illustrated in [40].) And so this configuration is associated with the permutation,

$$\sigma \equiv \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 7 & 6 & 10 & 9 & 8 & 13 & 12 \end{pmatrix}. \tag{5.7}$$

The definition of σ can be restated in an equivalent, but more transparently geometric form:

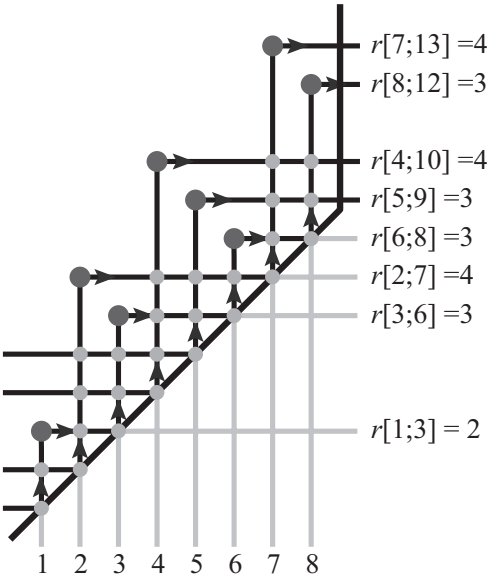
Definition: For each $a \in \{1,\ldots,n\}$, the permutation $\sigma(a) \geq a$ labels the *first* column $c_{\sigma(a)}$ such that $c_a \in \text{span}\{c_{a+1},\ldots,c_{\sigma(a)}\}$.

(Notice that if $c_a = \vec{0}$, then $\sigma(a) = a$, as $\vec{0}$ is spanned by the *empty* chain '$\{c_{a+1},\ldots,c_a\}$'.)

This definition is useful in practice. For example, it makes it easy to understand how the dimensionality of a configuration is encoded by its permutation. Notice that because $c_a \in \text{span}\{c_{a+1},\ldots,c_{\sigma(a)}\}$, we may expand c_a into the $r[a;\sigma(a)]$-dimensional space spanned by $\{c_{a+1},\ldots,c_{\sigma(a)}\}$; therefore, specifying c_a requires $r[a;\sigma(a)]$ degrees of freedom. And so, remembering to subtract the k^2 degrees of freedom absorbed by the overall $GL(k)$-redundancy, we find that:

$$\dim(C_\sigma) = \left(\sum_{a=1}^{n} r[a;\sigma(a)]\right) - k^2. \tag{5.8}$$

Notice that $r[a;\sigma(a)]$ is just exactly the number of other hooks that intersect the vertical (or horizontal) part of any particular hook $a \mapsto \sigma(a)$. Thus, for our example in $G(4,8)$ given above, the ranks $r[a,\sigma(a)]$ can be read off as the number of intersections (marked in green) along each vertical (or horizontal) line:

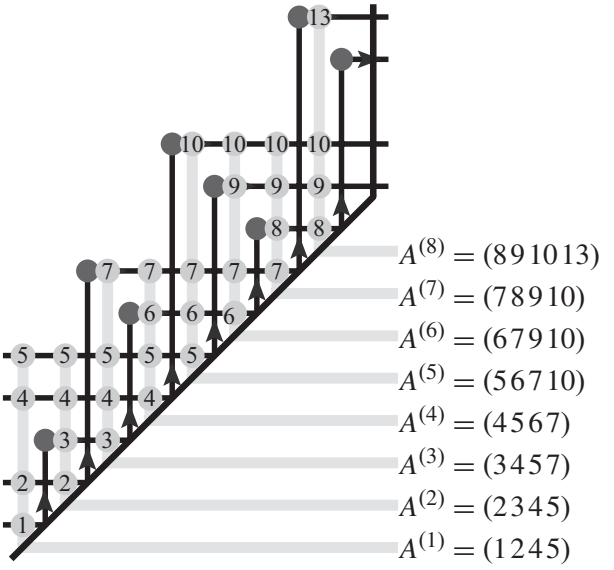


(5.9)

which shows that this configuration has $25-4^2=9$ degrees of freedom.

It is not hard to see how the permutation encodes *all* the ranks $r[a;b]$, thereby demonstrating that σ *fully* characterizes any configuration in the positroid stratification. If we let $q[a;b]$ denote the number of $c \in \{b-n,\ldots,a\}$ such that $\sigma(C) \in \{b,\ldots,a+n\}$, then $r[a;b] = k-q[a;b]$. Graphically, $q[a;b]$ is the number of hooks whose corners are above and to the left of $r[a;b]$ in the table (5.5).

The permutation is the most compact, most invariant way of describing the consecutive linear dependencies of a configuration of vectors. A more redundant, but sometimes useful alternative characterization of a configuration is known as the *Grassmannian necklace* [39]: a list of n, k-tuples $A^{(a)} \equiv (A_1^{(a)},\ldots,A_k^{(a)})$ denoting the lexicographically minimal non-vanishing minors starting from each of the n columns. Geometrically, $A^{(a)}$ encodes the labels of the *first* k column vectors beyond (or possibly including) c_a, for which $\text{rank}\{c_{A_1^{(a)}}, \ldots, c_{A_k^{(a)}}\} = k$. In terms of the hooks described above, $A^{(a)}$ simply lists the k horizontal lines that pass above the a th column (which often do not cross the hook going from $a \mapsto \sigma(a)$). In the $G(4,8)$ example above, (5.3), the Grassmannian necklace can be read off as follows:

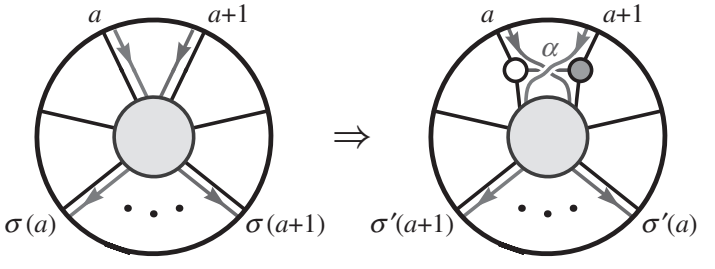


(5.10)

5.2 Canonical coordinates and the equivalence of permutation labeling

In Chapter 4, we saw that every on-shell graph is associated with both a permutation (via left-right paths) and also a k-plane in n dimensions $C \in G(k,n)$ encoding the linear relations involving the external data. And we have just seen that any such plane C, viewed as a configuration of column vectors, can *also* be labeled by a permutation. We will now demonstrate that these permutation labels match—that the configuration $C \in G(k,n)$ associated with an on-shell graph labeled by the left-right-path permutation σ, is labeled *geometrically* by the *same* permutation σ.

The proof of the equivalence of these permutation labels is both simple and constructive. Recall from section 3.2 that a representative reduced on-shell graph can be constructed for any permutation σ by decomposing it into a sequence of 'adjacent' transpositions acting on a trivial permutation, where each successive transposition in the decomposition adds a BCFW-bridge to the graph according to the following:



(5.11)

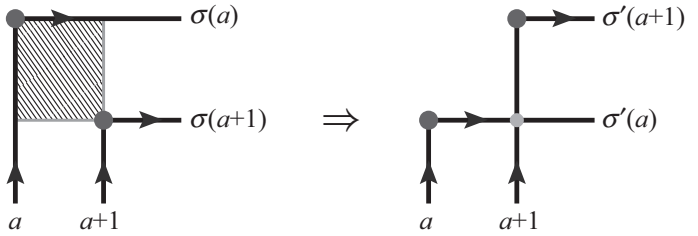
(As before, recall that two columns are to be considered 'adjacent' if separated only by columns that are self-identified under σ.) Now, just as we can build up a representative on-shell graph in this way for any permutation, we can also build up a representative matrix $C_\sigma \in G(k,n)$, which we will find to be labeled *geometrically* by the same permutation. As a bonus, this construction will provide us with explicit coordinates for any cell of the positroid, and these coordinates will have many nice properties.

What action on the columns of C corresponds to adding a BCFW-bridge, as in (5.11)? In terms of the matrices associated with on-shell diagrams, adding a bridge shifts:

$$c_{a+1} \mapsto c_{\widehat{a+1}} \equiv c_{a+1} + \alpha\, c_a; \tag{5.12}$$

recall also this shift changes the measure on the Grassmannian by adding a factor of $d\log(\alpha)$.

Notice that if we take a residue about $\alpha = 0$, we restore the original configuration; thus, $\alpha \mapsto 0$ can viewed as deleting the new edge from the graph in (5.11). Of course, in terms of the left-right path permutations, the BCFW-bridge transposes the images of a and $a+1$ under σ. What we need to show, therefore, is that the shift (5.12) has this same effect on the geometric permutation defined by the columns:



(5.13)

Let us now show that this is indeed the change induced by (5.12). Clearly, the transformation (5.12) can at most affect the ranks of chains that *include* c_{a+1} and *not* c_a. After the shift, $c_{\widehat{a+1}}$ is no longer spanned by $\{c_{a+2} \ldots, c_{\sigma(a+1)}\}$, because c_a is not; but $c_{\widehat{a+1}}$ *is* spanned by $\{c_{a+2}, \ldots, c_{\sigma(a)}\}$; and so, $\sigma(a+1) \mapsto \sigma'(a+1) = \sigma(a)$. Similarly, after the shift c_a is trivially in the span of $\{c_{\widehat{a+1}}, \ldots, c_{\sigma(a+1)}\}$ as $\text{span}\{c_{\widehat{a+1}}, \ldots, c_{\sigma(a+1)}\} = \text{span}\{c_a, c_{a+1}, \ldots, c_{\sigma(a+1)}\}$; and so, $\sigma(a) \mapsto \sigma'(a) = \sigma(a+1)$. And we are done.

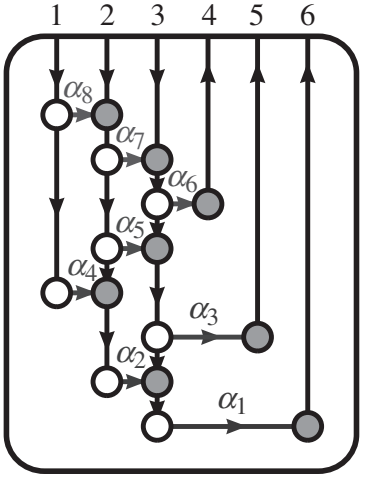
Therefore, just as successive BCFW-bridges, (5.11), can be used to construct a representative, reduced on-shell graph for any permutation, they also provide us with a representative matrix for the configuration—and the BCFW-shift parameters, denoted α_i, provide us with coordinates.

We can see how this works explicitly by revisiting the example given in section 3.2 where we used successive BCFW-bridges to construct a representative on-shell graph for the permutation $\{4,6,5,7,8,9\}$. Repeating the same construction as before, but now decorating each BCFW-bridge with its corresponding

shift-parameter α_i gives rise to the following:

τ	1 2 3 4 5 6 ↓ ↓ ↓ ↓ ↓ ↓	BCFW shift
	4 6 5 7 8 9	
(1 2)		$c_2 \mapsto c_2 + \alpha_8 c_1$
	6 4 5 7 8 9	
(2 3)		$c_3 \mapsto c_3 + \alpha_7 c_2$
	6 5 4 7 8 9	
(3 4)		$c_4 \mapsto c_4 + \alpha_6 c_3$
	6 5 7 4 8 9	
(2 3)		$c_3 \mapsto c_3 + \alpha_5 c_2$
	6 7 5 4 8 9	
(1 2)		$c_2 \mapsto c_2 + \alpha_4 c_1$
	7 6 5 4 8 9	
(3 5)		$c_5 \mapsto c_5 + \alpha_3 c_3$
	7 6 8 4 5 9	
(2 3)		$c_3 \mapsto c_3 + \alpha_2 c_2$
	7 8 6 4 5 9	
(3 6)		$c_6 \mapsto c_6 + \alpha_1 c_3$
	7 8 9 4 5 6	

$\Leftrightarrow$



Starting with the zero-dimensional configuration labeled by {7,8,9,4,5,6} and performing each successive BCFW-shift generates the following representation of C:

$$C(\vec{\alpha}) \equiv \begin{pmatrix} 1 & (\alpha_4+\alpha_8) & \alpha_4(\alpha_5+\alpha_7) & \alpha_4\alpha_5\alpha_6 & 0 & 0 \\ 0 & 1 & (\alpha_2+\alpha_5+\alpha_7) & (\alpha_2+\alpha_5)\alpha_6 & \alpha_2\alpha_3 & 0 \\ 0 & 0 & 1 & \alpha_6 & \alpha_3 & \alpha_1 \end{pmatrix}. \qquad (5.14)$$

For the sake of illustration, below is the complete sequence of coordinatized cells generated along the chain of BCFW-shifts that build up $C(\alpha)$:

$$\begin{pmatrix} 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&1&0&0&0 \end{pmatrix}_{\{7,8,\mathbf{9},4,5,\mathbf{6}\}} \xrightarrow[\alpha_1]{(36)} \begin{pmatrix} 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&1&0&0&\alpha_1 \end{pmatrix}_{\{7,\mathbf{8},\mathbf{6},4,5,9\}} \xrightarrow[\alpha_2]{(23)} \begin{pmatrix} 1&0&0&0&0&0 \\ 0&1&\alpha_2&0&0&0 \\ 0&0&1&0&0&\alpha_1 \end{pmatrix}_{\{7,6,\mathbf{8},4,\mathbf{5},9\}} \xrightarrow[\alpha_3]{(35)} \begin{pmatrix} 1&0&0&0&0&0 \\ 0&1&\alpha_2&0&\alpha_2\alpha_3&0 \\ 0&0&1&0&\alpha_3&\alpha_1 \end{pmatrix}_{\{\mathbf{7},\mathbf{6},5,4,8,9\}}$$

$$\downarrow (12)\,\alpha_4$$

$$\begin{pmatrix} 1&\alpha_4&\alpha_4\alpha_5&\alpha_4\alpha_5\alpha_6&0&0 \\ 0&1&(\alpha_2+\alpha_5)&(\alpha_2+\alpha_5)\alpha_6&\alpha_2\alpha_3&0 \\ 0&0&1&\alpha_6&\alpha_3&\alpha_1 \end{pmatrix}_{\{6,\mathbf{5},\mathbf{4},7,8,9\}} \xleftarrow[\alpha_6]{(34)} \begin{pmatrix} 1&\alpha_4&\alpha_4\alpha_5&0&0&0 \\ 0&1&(\alpha_2+\alpha_5)&0&\alpha_2\alpha_3&0 \\ 0&0&1&0&\alpha_3&\alpha_1 \end{pmatrix}_{\{6,5,\mathbf{7},\mathbf{4},8,9\}} \xleftarrow[\alpha_5]{(23)} \begin{pmatrix} 1&\alpha_4&0&0&0&0 \\ 0&1&\alpha_2&0&\alpha_2\alpha_3&0 \\ 0&0&1&0&\alpha_3&\alpha_1 \end{pmatrix}_{\{6,\mathbf{7},\mathbf{5},4,8,9\}}$$

$$\downarrow (23)\,\alpha_7$$

$$\begin{pmatrix} 1&\alpha_4&\alpha_4(\alpha_5+\alpha_7)&\alpha_4\alpha_5\alpha_6&0&0 \\ 0&1&(\alpha_2+\alpha_5+\alpha_7)&(\alpha_2+\alpha_5)\alpha_6&\alpha_2\alpha_3&0 \\ 0&0&1&\alpha_6&\alpha_3&\alpha_1 \end{pmatrix}_{\{\mathbf{6},\mathbf{4},5,7,8,9\}} \xrightarrow[\alpha_8]{(12)} \begin{pmatrix} 1&(\alpha_4+\alpha_8)&\alpha_4(\alpha_5+\alpha_7)&\alpha_4\alpha_5\alpha_6&0&0 \\ 0&1&(\alpha_2+\alpha_5+\alpha_7)&(\alpha_2+\alpha_5)\alpha_6&\alpha_2\alpha_3&0 \\ 0&0&1&\alpha_6&\alpha_3&\alpha_1 \end{pmatrix}_{\{4,6,5,7,8,9\}}$$

Coordinates generated in this way enjoy many nice properties. For example, the physically relevant measure on the Grassmannian (integration over which generates the on-shell differential forms of interest) is *maximally* simple in these coordinates: because each BCFW-shift simply adds a factor of $d\log(\alpha)$ to the measure, the final measure is simply,

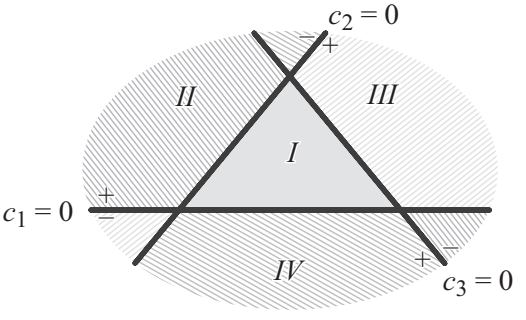
$$\frac{d\alpha_1}{\alpha_1} \wedge \cdots \wedge \frac{d\alpha_d}{\alpha_d} = d\log(\alpha_1) \wedge \cdots \wedge d\log(\alpha_d)\,. \tag{5.15}$$

Another important property—to be described more fully in section 6.4—is that these coordinates make it possible to access each of the lower-dimensional boundaries of C as the zero-loci of some of the α_i (using an atlas of at most n coordinate charts).

5.3 Positroid cells and the positive part of the Grassmannian

So far in our discussion of configurations of vectors we have only discussed basic, linear dependencies. Let us now consider the case where these vectors are real. This will expose a natural and beautiful object, known as the *positive Grassmannian*, denoted $G_+(k,n)$. As in the previous subsection, let us first jump ahead and describe this object intrinsically, and then return to on-shell diagrams and show how the amalgamation picture described in section 4.4 makes it obvious that on-shell diagrams—whether reduced or not—are always associated with points in $G_+(k,n)$, and demonstrate how this works explicitly for the reduced graphs obtained via the BCFW-bridge decomposition described in the previous section.

Perhaps the best way to motivate the positive Grassmannian is by starting with the simplest case, $G_{\mathbb{R}}(1,n) \simeq \mathbb{RP}^{n-1}$. Here, the column 'vectors' c_a of a 1-plane $C \equiv (c_1, \ldots, c_n)$ are simply homogeneous coordinates on $\mathbb{RP}^{n-1}$, and the 'positive part' of $\mathbb{RP}^{n-1}$ is simply the part of projective space where all the homogeneous coordinates are positive, which is nothing but a simplex. Consider for example $\mathbb{RP}^2$ corresponding to the 1-plane $C \equiv (c_1, c_2, c_3)$:



(5.16)

The 'positive part' of $\mathbb{RP}^2$ is defined by the region where all the homogeneous coordinates c_a are positive—corresponding to the (open) region labeled "I" above. Of course, because we often allow ourselves to rescale each $c_a \sim t_a c_a$, *any* relative signs among the homogeneous coordinates will describe an open region of $\mathbb{RP}^2$ essentially equivalent to region I, dividing $\mathbb{RP}^2$ into four "positive parts" as indicated in (5.16). Continuing this logic to higher n, it is clear that the "positive part" of $\mathbb{RP}^{n-1}$ should be defined as the (open) simplex for which all homogeneous coordinates are positive.

For higher k, the "positive part" of $G(k,n)$ is a natural generalization of the notion of a simplex in $G(1,n)$. Thinking of the homogeneous coordinates c_a as (1×1) 'minors' of $C\in G(1,n)$, it is natural to define the positive part of $G(k,n)$ to be the region for which all *ordered* minors $(a_1\cdots a_k)$, with $a_1<\cdots<a_k$, are positive. (Notice that without a fixed ordering of the columns, it would be meaningless to discuss the positivity of minors as they are antisymmetric with respect to ordering.)

Although this definition of the positive part of $G(k,n)$ requires an ordering of the columns, no reference was made to any *cyclic* structure. But cyclicity emerges automatically. Naïvely, it would seem that there could be a distinct positive part for each of the $n!$ orderings of the columns, but some of these are actually the same. Suppose that $C\in G_+(k,n)$ for columns ordered according to $\{c_1,\ldots,c_n\}$. Then the change

$$c_1\to c_2,\, c_2\to c_3,\,\cdots,\, c_n\to(-1)^{k+1}c_1, \tag{5.17}$$

gives a positive configuration in the rotated ordering. This is referred to as a "twisted" cyclic symmetry.

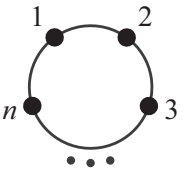
Notice that the definition of $G_+(k,n)$ has so far made no reference to *consecutivity* of the constraints involved in its boundary configurations (where some minors are allowed to vanish). The reason why consecutivity plays a role is that not all minors are independent—recall from section 4.1 that they satisfy Plücker relations following from Cramer's rule, (4.4). The relevance of this will become clear in a simple example. Consider the case of $G(2,4)$, where we have

$$(13)(24)=(12)(34)+(14)(23). \tag{5.18}$$

Notice the presence of the *plus* sign on the right-hand side. It implies that if we start with a configuration in $G_+(2,4)$, the minor (13) can only vanish if at least two other ordered minors also vanish.

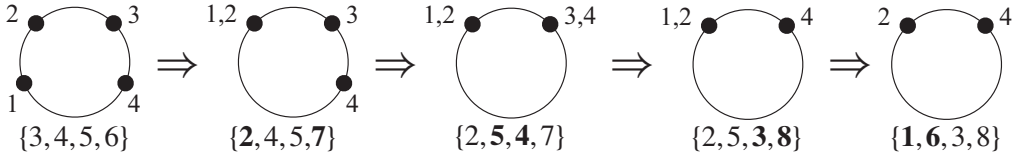
We can see how consecutivity matters more generally for $G(2,n)$ by thinking of the column vectors projectively as points in $\mathbb{RP}^1$. If we rescale the columns to be of the form $c_a\sim\left(\begin{smallmatrix}\beta_a\\1\end{smallmatrix}\right)$, then $(ab)=(\beta_a-\beta_b)$, and so a positive configuration is simply one for which $\beta_a>\beta_b$ for all $a<b$. That is, the positive part of $G(2,n)$ is

just precisely the configurations of *ordered* points on a circle:



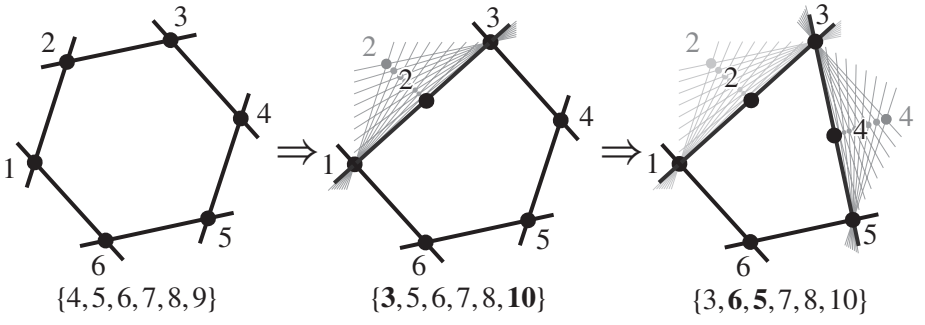
(5.19)

As such, it is clear that codimension-one boundaries should correspond to the vanishing of only *consecutive* minors—the collision of adjacent points in $\mathbb{RP}^1$. In $G(2,4)$, for example, the following sequence of boundaries connect a generic configuration to one without any degrees of freedom:



In order to see that this phenomenon is not peculiar to $G(2,n)$, and to get a better picture for what is going on, let us look again at $G(3,n)$. We may use the rescaling symmetry to write each column as $c_a \sim \begin{pmatrix} \widehat{c}_a \\ 1 \end{pmatrix}$, where each $\widehat{c}_a$ is in $\mathbb{R}^2$. It is then easy to check that the requirement of positivity for all ordered minors translates into the geometric statement that the points $\widehat{c}_a$ form the vertices of a *convex* polygon in the plane.

Because of convexity, it is easy to see that going to boundaries can only involve linear relations between *consecutive* chains of columns. For instance, below we draw a projective representation of a generic configuration $G(3,6)$, and some of the boundaries obtainable while preserving convexity:



From the generic configuration, it is possible to make any consecutive minor vanish such as (1 2 3) shown above. Projectively, a minor will vanish whenever three points become collinear. However, note that for instance the non-consecutive minor (1 3 5) cannot be made to vanish without either: 1. destroying convexity, or 2. forcing *additional* minors to vanish along the way. And so, we find the

same stratification of successive boundaries as those obtained by consecutive constraints.

These examples suffice to motivate a remarkable connection, which we will shortly understand in a simple and general way. In the first part of this section, we discussed a stratification of the *complex* Grassmannian, in terms of specified linear dependencies between consecutive column vectors. We now see that this structure is beautifully characterized by the structure of the *real* Grassmannian: the cell decomposition of the positive Grassmannian is precisely specified by giving linear dependencies between consecutive vectors.

But first, let us step back and understand the simple and direct connection between on-shell diagrams and the positive Grassmannian. Recall that we can construct the configuration $C \in G(k,n)$ for any on-shell diagram by simply "amalgamating" the 1- and 2-planes associated with the white and the black vertices, respectively. We saw in section 4.4 that only two operations were needed to construct the plane $C \in G(k,n)$ for any on-shell graph: combining graphs via *direct products*, and gluing legs together by *projecting out* on-shell pairs of particles. Let us briefly recall how these two operations act on the minors of the planes involved, and verify the wonderful fact that *amalgamation preserves positivity*.

The proof is simple. First, observe that we can always use rescaling symmetry to make any configuration in $G(1,3)$ or $G(2,3)$ positive (see, e.g. (5.16)). Therefore, an on-shell graph can always be constructed by attaching these positive cells to each vertex, and then proceeding with amalgamation as described in section 4.4. Recall that the simplest of the two operations, taking direct products, acts trivially on minors: suppose that the columns of $C_L \in G(k_L,n_L)$ are ordered $\{c_1,\ldots,c_{n_L}\}$, and that those of $C_R \in G(k_R,n_R)$ are ordered $\{c_{n_L+1},\ldots,c_{n_L+n_R}\}$, then all the non-vanishing minors $C_L \bigotimes C_R \mapsto C \in G(k_L+k_R,n_L+n_R)$ will be given by

$$(a_1 \cdots a_{k_L} b_1 \cdots b_{k_R})\big|_C = (a_1 \cdots a_{k_L})\big|_{C_L} \times (b_1 \cdots b_{k_R})\big|_{C_R}\,; \tag{5.20}$$

and so, if C_L and C_R are both *positive*, then C will be as well.

The second fundamental operation, called *projection*, takes a configuration $C \in G(k+1,n+2)$ and produces a configuration $\widehat{C} \in G(k,n)$, obtained by projecting C into the orthogonal complement of $(c_A - c_B)$, for two *adjacent* legs (AB). In terms of minors, this operation acts according to:

$$(a_1 \cdots a_k)|_{\widehat{C}} = (A\, a_1 \cdots a_k)|_C + (B\, a_1 \cdots a_k)|_C\,. \tag{5.21}$$

If (AB) are the *first two* labels for the columns of $C \in G_+(k+1,n+2)$, then both terms on the right hand side are trivially positive; if (AB) are not the first two columns, then they can always be brought to this position at the trivial cost of rescaling some columns by (-1) as described during our discussion of the twisted cyclic structure of $G_+(k,n)$ in section 5.3.

5.4 Canonically positive coordinates for positroids

We have seen many ways to describe the configuration $C \in G(k,n)$ associated with an on-shell diagram, including procedures that explicitly generate a matrix representative of C parameterized by variables attached to the *faces* or the *edges* of a graph (see section 4.5). And in section 5.2, we saw that "canonical" coordinates for any cell $C \in G(k,n)$ in the positroid stratification can be systematically generated (along with a representative, reduced graph) by applying successive BCFW-shifts. In this subsection, we demonstrate that a slight-modification of these BCFW-bridge coordinates (see equation (5.27)) have the remarkable property that when the coordinates α_i are themselves positive, then $C(\alpha_i) \in G_+(k,n)$! We will refer to any such coordinates that have this property as "positive."

Before we describe how the BCFW-bridge coordinates make positivity manifest in this way, let us first describe a more intuitive way to parameterize generic configurations in $G(k,n)$ with coordinates that share this property. It will turn out that this geometrically motivated parameterization of $G(k,n)$ will be essentially identical to that which is generated by the BCFW-bridge construction, and so this slight detour will prove itself quite useful later (see Chapter 7).

Observe that any homogeneous coordinates for $G(1,n) \simeq \mathbb{P}^{n-1}$ are trivially positive:

$$C^{(1,n)} \equiv \begin{pmatrix} \beta_1^1 & \beta_2^1 & \cdots & \beta_{n-1}^1 & \beta_n^1 \end{pmatrix}, \tag{5.22}$$

because $C^{(1,n)}(\beta) \in G_+(1,n)$ whenever all the variables $\beta_a^1 > 0$.

The first nontrivial case is for $G(2,n)$. Recall from our discussion above that if we rescale all the column vectors of $C \in G(2,n)$ to be of the form $c_a \sim \begin{pmatrix} \widehat{c}_a \\ 1 \end{pmatrix}$, then $(ab) = \widehat{c}_a - \widehat{c}_b$; and so any set of ordered numbers $\widehat{c}_1 > \cdots > \widehat{c}_n$ will parameterize a point in $G_+(2,n)$. One natural way to create such an ordered list of positive numbers would be to have $\widehat{c}_a = \widehat{c}_{a+1} + \beta_{a+1}^1$ for arbitrary, positive β_{a+1}^1—where we have intentionally named these 'arbitrary' positive parameters according to our parameterization of $G_+(1,n)$ in (5.22). Restoring the degrees of freedom that rescale each column vector, we obtain the following:

$$C^{(2,n)} \equiv \begin{pmatrix} \beta_1^2(\beta_2^1 + \cdots) & \beta_2^2(\beta_3^1 + \cdots) & \cdots & \beta_{n-1}^2(\beta_n^1) & 0 \\ \beta_1^2 & \beta_2^2 & \cdots & \beta_{n-1}^2 & \beta_n^2 \end{pmatrix}. \tag{5.23}$$

It is easy to verify that if $\beta_a^\alpha > 0$, then $C^{(2,n)}(\beta) \in G_+(2,n)$.

This construction naturally continues recursively, generating positive coordinates for any (generic) configuration in $G(k,n)$ as follows:

$$C^{(k,n)} \equiv \begin{pmatrix} \beta_1^k \widehat{c}_1^{(k,n)} & \cdots & \beta_{n-1}^k \widehat{c}_{n-1}^{(k,n)} & 0 \\ \beta_1^k & \cdots & \beta_{n-1}^k & \beta_n^k \end{pmatrix} \text{ with } \widehat{c}_a^{(k,n)} \equiv \sum_{j=(a+1)}^{n} c_j^{(k-1,n)}. \tag{5.24}$$

Surprisingly, after using $GL(k)$-redundancy to remove the excess degrees of freedom in the parameterization of $C^{(k,n)}(\beta)$, it turns out that these are (essentially) identical to the coordinates produced by the BCFW-bridge construction described in section 5.2. Indeed, the only distinction is a relabeling of bridge-variables $\alpha_1,\ldots,\alpha_d$ (where $d\equiv\dim(G(k,n))=k(n-k)$) according to:

$$\begin{pmatrix} \beta^1_{k+1} & \beta^1_{k+2} & \cdots & \beta^1_{n-1} & \beta^1_n \\ \beta^2_{k+1} & \beta^2_{k+2} & \cdots & \beta^2_{n-1} & \beta^2_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta^k_{k+1} & \beta^k_{k+2} & \cdots & \beta^k_{n-1} & \beta^k_n \end{pmatrix} \Leftrightarrow \begin{pmatrix} \alpha_d & \alpha_{d-2} & \cdots & & \cdots & \cdots & \cdots & \alpha_\ell & \cdots & \cdots \\ \alpha_{d-1} & \ddots & \ddots & & \ddots & \ddots & \alpha_{\ell+1} & \ddots & \ddots & \vdots \\ & \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & \ddots & \alpha_2 \\ & & \cdots & \cdots & \alpha_{\ell+k-1} & \cdots & \cdots & \cdots & \alpha_3 & \alpha_1 \end{pmatrix}$$

Let us now show that positivity is a *manifest* property of the BCFW-bridge coordinates for *all* positroid cells. This will also complete the connection between on-shell graphs, the stratification of configurations of vectors given by prescribing linear dependencies between consecutive vectors, and the cell decomposition of the *positive* Grassmannian.

We begin by observing that the minors of C transform nicely under BCFW-shifts:

$$(\cdots a+1\cdots)\mapsto(\cdots\widehat{a+1}\cdots)=(\cdots a+1\cdots)+\alpha\,(\cdots a\cdots). \tag{5.25}$$

And so, if we start with a configuration C in the positive Grassmannian, and if a and $a+1$ are *strictly* adjacent—with no columns between them self-identified under σ—then the BCFW shift preserves positivity, because whenever $(\cdots a+1\cdots)$ is ordered, so is $(\cdots a\cdots)$.

However, we must remember that the decomposition of a permutation into 'adjacent' transpositions allows for a and "$a+1$" to be separated by any number of columns that map to themselves (mod n) under σ. Because $\sigma(b)=b$ (as opposed to $\sigma(b)=b+n$) implies that $c_b=0$, all minors involving b vanish; and so, skipping-over these columns will not affect any non-vanishing minors. However, $\sigma(b)=b+n$ if and only if $c_b\notin\text{span}\{c_{b+1},\ldots,c_{b+n-1}\}$, implying that c_b is not spanned by the rest of the columns of C; as such, $\sigma(b)=b+n$ implies that b *must* be involved in any non-vanishing $(k\times k)$ minor of C. And so, when this happens, the shift in (5.25) may not preserve ordering for both of the terms.

To illustrate this minor subtlety, consider the very simplest case in which it arises: the one-dimensional configuration $C\in G(2,3)$ labeled by the permutation $\sigma\equiv\{3,5,4\}$. The decomposition of σ into 'adjacent' transpositions involves only one step: (1 3) —an 'adjacent' transposition that skips over column c_2 because $\sigma(2)=2+3$. Explicitly, the BCFW-coordinates of C_σ would be generated as

follows:

τ	$\begin{matrix}1&2&3\\\downarrow&\downarrow&\downarrow\end{matrix}$	BCFW shift
(1 3)	$\begin{matrix}3&5&4\\4&5&3\end{matrix}$	$c_3 \mapsto c_3 + \alpha_1 c_1$

$$\underset{\{\mathbf{4},5,\mathbf{3}\}}{\begin{pmatrix}1&0&0\\0&1&0\end{pmatrix}} \xrightarrow[\alpha_1]{(13)} \underset{\{3,5,4\}}{\begin{pmatrix}1&0&\alpha_1\\0&1&0\end{pmatrix}} \tag{5.26}$$

Notice that the minor (23), which vanishes before the shift, becomes $(23) \mapsto (\widehat{23}) = (23) + \alpha_1(21) = -\alpha_1(12)$ after the shift. And so, if we wish to make the final configuration *C positive*, we must take α_1 to be negative; alternatively, we could redefine the rule for BCFW-shifts so that the transposition (13) actually corresponds to a shift $c_3 \mapsto c_3 - \alpha_1 c_1$. Of the two alternatives, we prefer the latter as then positivity of the BCFW-shift coordinates would directly imply that a configuration were positive.

It is easy to see how this simple example generalizes: in order to preserve the positivity of minors *and* the coordinates, we should redefine the BCFW-shift so that the transposition of a and "$a+1$" changes the columns of C according to

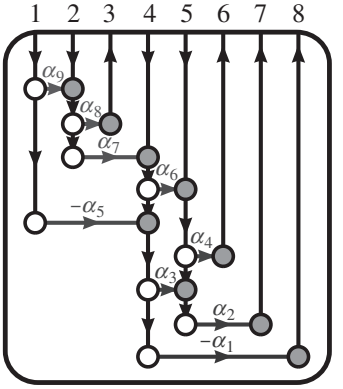
$$c_{a+1} \mapsto c_{a+1} + (-1)^q \alpha\, c_a, \tag{5.27}$$

where q is the number of columns b between a and "$a+1$" such that $\sigma(b) = b + n$. In this modified form, the BCFW-shift is *guaranteed* to preserve positivity. And so, restricting all the coordinates α_i to be positive will always result in a configuration $C(\vec{\alpha})$ in the *positive* Grassmannian $G_+(k,n)$.

To see how these *signed* BCFW-shifts make positivity manifest—and as one further example of the BCFW-bridge construction described in section 5.2—consider the following coordinates constructed for the configuration in $G(4,8)$ given in (5.3), labeled by the permutation $\{3,7,6,10,9,8,13,12\}$:

τ	q	$\begin{matrix}1&2&3&4&5&6&7&8\\\downarrow&\downarrow&\downarrow&\downarrow&\downarrow&\downarrow&\downarrow&\downarrow\end{matrix}$	BCFW shift
(1 2)	0	3 7 6 10 9 8 13 12	$c_2 \mapsto c_2 + \alpha_9 c_1$
(2 3)	0	7 3 6 10 9 8 13 12	$c_3 \mapsto c_3 + \alpha_8 c_2$
(2 4)	0	7 6 3 10 9 8 13 12	$c_4 \mapsto c_4 + \alpha_7 c_2$
(4 5)	0	7 10 3 6 9 8 13 12	$c_5 \mapsto c_5 + \alpha_6 c_4$
(1 4)	1	7 10 3 9 6 8 13 12	$c_4 \mapsto c_4 - \alpha_5 c_1$
(5 6)	0	9 10 3 7 6 8 13 12	$c_6 \mapsto c_6 + \alpha_4 c_5$
(4 5)	0	9 10 3 7 8 6 13 12	$c_5 \mapsto c_5 + \alpha_3 c_4$
(5 7)	0	9 10 3 8 7 6 13 12	$c_7 \mapsto c_7 + \alpha_2 c_5$
(4 8)	1	9 10 3 8 13 6 7 12	$c_8 \mapsto c_8 - \alpha_1 c_4$
		9 10 3 12 13 6 7 8	

$\Leftrightarrow$



(5.28)

As with the example (5.14) given above, we can construct the corresponding chart $C(\alpha)$ according to the successive BCFW shifts:

$$C(\alpha) \equiv \begin{pmatrix} 1 & \alpha_9 & 0 & -\alpha_5 & -\alpha_5\alpha_6 & 0 & 0 & 0 \\ 0 & 1 & \alpha_8 & \alpha_7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \alpha_3+\alpha_6 & \alpha_3\alpha_4 & 0 & -\alpha_1 \\ 0 & 0 & 0 & 0 & 1 & \alpha_4 & \alpha_2 & 0 \end{pmatrix} \tag{5.29}$$

It is easy to verify that all the non-vanishing minors of $C(\alpha) \in G(4,8)$ are positive when $\alpha_i \in \mathbb{R}_+$. For example, consider the minor

$$(2457) = \alpha_2\alpha_3\alpha_5 + \alpha_2\alpha_3\alpha_7\alpha_9 + \alpha_2\alpha_6\alpha_7\alpha_9, \tag{5.30}$$

the positivity of which requires, for example, the *signed* BCFW-shift $c_4 \mapsto c_4 - \alpha_5 c_1$.

6

Boundary configurations, graphs, and permutations

6.1 Physical singularities and positroid boundaries

Recall that an on-shell diagram labeled by the permutation σ corresponds to a differential form f_σ obtained via integration over the configuration $C_\sigma(\alpha) \in G(k,n)$ subject to the constraints that C_σ be orthogonal to $\widetilde{\lambda}$ and contain λ:

$$f_\sigma = \int_{C_\sigma} \frac{d\alpha_1}{\alpha_1} \wedge \cdots \wedge \frac{d\alpha_d}{\alpha_d}\, \delta^{k\times 4}\big(C_\sigma \cdot \widetilde{\eta}\big)\delta^{k\times 2}\big(C_\sigma \cdot \widetilde{\lambda}\big)\delta^{2\times(n-k)}\big(\lambda \cdot C_\sigma^{\perp}\big), \tag{6.1}$$

where α_i are *canonical* (e.g. BCFW-bridge) coordinates for the configuration C_σ. Because the δ-functions encode $(2n-4)$ constraints in general (together with the four constraints of momentum-conservation), cells with $(2n-4)$ degrees of freedom can be fully localized, while those of lower dimension leave behind further δ-functions that impose constraints on the external kinematical data.

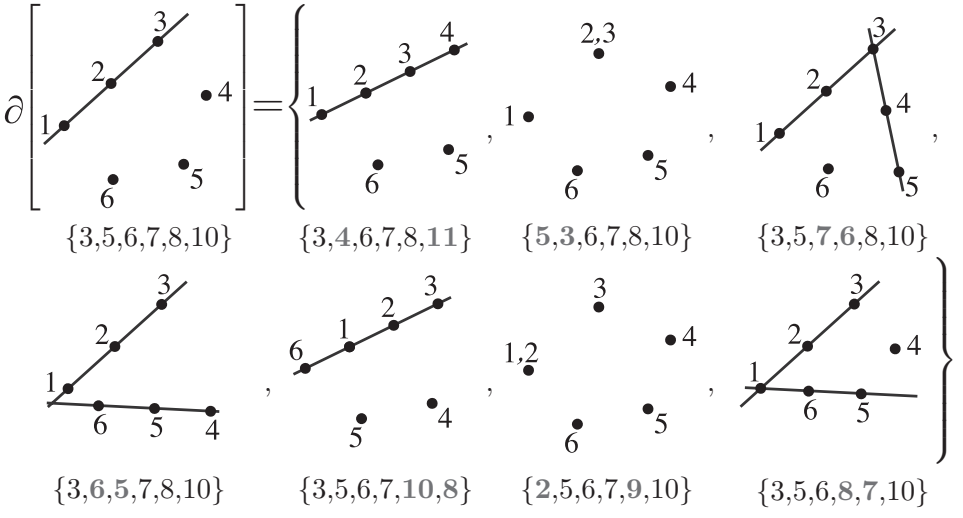
On-shell differential forms that impose constraints on the external data (beyond momentum conservation) represent physical *singularities*: places in the space of kinematical data where higher-degree forms develop poles. As we saw in section 2.6, such singularities are of primary physical interest: for example, knowing the singularity structure of scattering amplitudes suffices to fix them completely to all loop-orders via the BCFW recursion relations (2.54).

The physical singularities of on-shell differential forms, therefore, correspond to the *boundaries* of the corresponding configurations in the Grassmannian. Suppose we consider a *reduced* graph with n_F faces; then, because such a graph is associated with an (n_F-1)-dimensional configuration C, it is easy to see that its boundaries are those graphs obtained by deleting edges (reducing the number of faces by one). However, sometimes a graph obtained in this way is no longer reduced, and actually corresponds to a configuration in the Grassmannian whose dimension has been lowered by more than one. This raises the question: which edges in a graph can be removed while keeping a graph reduced? Such edges will be called *removable*. It turns out that this question is easiest to answer

not in terms of on-shell graphs directly, but in terms of the geometry of their corresponding configurations in the Grassmannian and the combinatorics of their permutations.

6.2 Boundary configurations: combinatorics and stratification

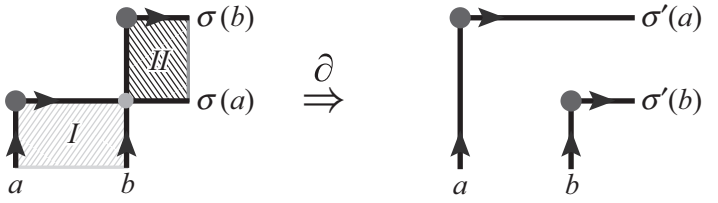
The codimension 1 boundaries of a configuration C, denoted $\partial(C)$, in the positroid stratification are those configurations obtained by imposing any one additional constraint involving consecutive chains of columns. Before describing the combinatorial rule for finding boundary configurations, let us first build some intuition through simple examples. Recall from section 5.1 the configuration in $G_+(3,6)$ whose boundaries were:



where we have highlighted how the permutation changes for each boundary-element.

And so—if it weren't sufficiently obvious already—this example makes it clear that boundary elements of a configuration labeled by σ are those labeled by σ' that are related to σ by a transposition of its images. However, not all transpositions lower the dimension of the configuration, and some transpositions lower the dimensionality by more than one. The way to identify the transpositions that lower the dimension by precisely one is easily understood from the way dimensionality is encoded by a configuration's permutation: if we view the permutation as given by the 'hooks' described in section 5.1, then the dimension of a configuration is counted by the number of intersections of its hooks (minus k^2). Therefore, boundaries are those transpositions that eliminate any *one* such

intersection:

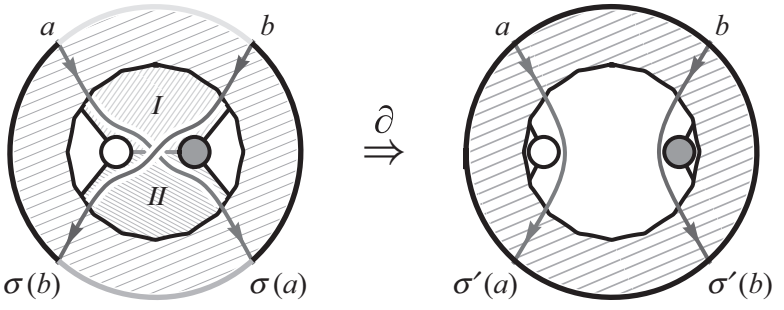


(6.2)

Here, it is important that $a < b \leq \sigma(a) < \sigma(b) \leq (a+n)$, and that there are no hooks from $c \in I$ to $\sigma(c) \in II$ as otherwise the dimensionality would be lowered by more than one:

(6.3)

Restated in terms of on-shell graphs decorated by left-right paths, this rule identifies *removable* edges as those along which two paths cross, $a \to \sigma(a)$ and $b \to \sigma(b)$ with $a < b \leq \sigma(b) < \sigma(a) \leq (a+n)$, provided that there is no path $c \to \sigma(c)$ with $c \in I$ and $\sigma(c) \in II$:



(6.4)

These two definitions of the boundary elements of a configuration are of course equivalent; but without the combinatorial rule for counting dimensions, it would have been considerably more difficult to see that these—and only these—edges are removable.

6.3 (Combinatorial) polytopes in the Grassmannian

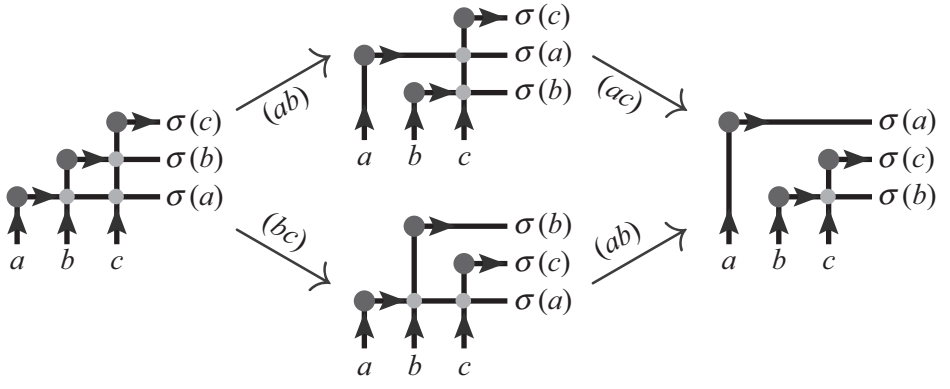
The boundary operator ∂ given on the previous page *defines* the positroid stratification of $G(k,n)$; and this stratification is a very special one, with many nice features. For one thing, it allows us to view every positroid configuration in $G_+(k,n)$ as *something like* a 'polytope' in $G(k,n)$. By this we mean that the inclusions induced by ∂ (viewed as a strong Bruhat covering relation) define an

Eulerian poset—the key combinatorial property of the poset of faces of an ordinary polytope.

We will not prove that ∂ defines an Eulerian poset (this was proven in [105]), but let us at least demonstrate that $\partial^2 = 0 \pmod 2$—which is of course a prerequisite for ∂ to actually have the meaning of a homological 'boundary' operator. It turns out that every configuration in $\partial^2(C)$ is found as the boundary of precisely *two* configurations in $\partial(C)$ (a fact that follows trivially from the more complete statement that ∂ defines an Eulerian poset). This is not hard to prove, and it trivially implies that $\partial^2 = 0 \pmod 2$. To see this, notice that each configuration in $\partial(C_\sigma)$ is labeled by σ' related to σ by a transposition. It is easy to see that the pair of transpositions must involve at least three distinct labels. If the pair involved four labels, say (ab) and (cd), then obviously the two transpositions can be taken in either order. When the pair involves three labels, say $\{abc\}$, then there are only four possible scenarios to check:

$$\begin{array}{ll} (ab)\circ(ac) \simeq (bc)\circ(ab) & (ab)\circ(bc) \simeq (bc)\circ(ac) \\ (ab)\circ(bc) \simeq (ac)\circ(ab) & (ac)\circ(bc) \simeq (bc)\circ(ab) \end{array} ; \qquad (6.5)$$

the first of these, for example, can be understood graphically in terms of hooks as,

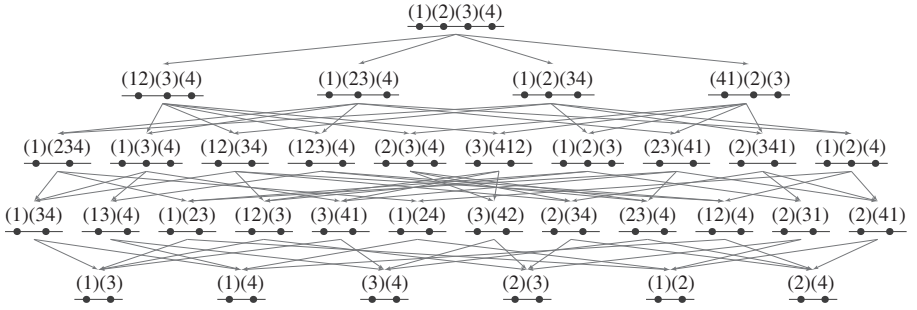


A more immediate, but somewhat indirect proof of this fact follows from the association of each permutation σ with a reduced on-shell graph. Recall that the graphs in the boundary of an on-shell graph labeled by σ are those for which *one* edge has been removed. Because each pair of left-right paths $a \to \sigma(a)$ and $b \to \sigma(b)$ cross on at most one edge of any reduced graph (if the edge is removable), it is clear that graphs in ∂^2 are those obtained by removing a pair of edges. As such, the pair of edges can be removed in any order, proving that there are two paths from any graph to each graph in ∂^2.

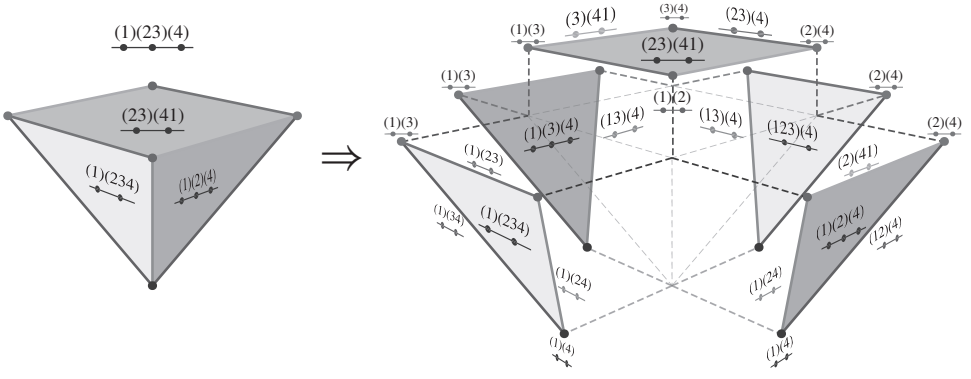
As mentioned above, an amazing feature of the positroid stratification is that the combinatorial structure of the inclusions induced by ∂ have the property that every positroid configuration defines an Eulerian poset—a combinatorial polytope.

Because of this, we can loosely view each positroid configuration as a region of $G(k,n)$ with essentially the topology of an open ball—even though such a picture is only strictly known to be valid for relatively simple cases such as $G(2,n)$.

In the case of the positroid $G_+(2,4)$, the polytope is relatively easy to visualize. The four-dimensional top-cell has four three-dimensional boundary configurations; the boundaries of these cells collectively involve ten two-dimensional configurations; and so on Starting with the generic configuration in $G_+(2,4)$, we find the boundaries defined by ∂ given as follows [106]:



Although it is hard to draw the complete four-dimensional polytope, its four three-dimensional faces each define square-pyramidal regions of $G(2,4)$. For example, the polytope corresponding to the configuration $\overset{(1)(23)(4)}{\bullet\!\!-\!\!\bullet\!\!-\!\!\bullet}$ of $G(2,4)$ labeled by the permutation $\{4,3,5,6\}$ is arranged as follows:



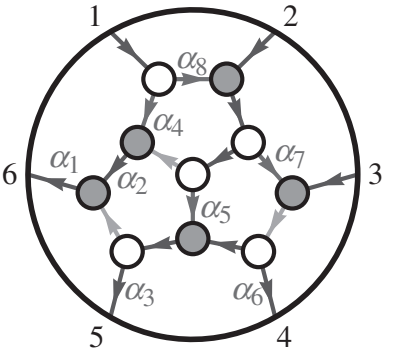
6.4 Approaching boundaries in canonical coordinates

Recall that the singularities of an on-shell differential form associated with an on-shell diagram are simply the residues of its poles. When written in terms of canonical coordinates on the Grassmannian as described above (see equation (6.1)), it is tempting to identify the manifestly logarithmic singularities in the measure with configurations in the 'boundary.' But there are two important points

that make such a correspondence a bit more delicate than it may appear at first glance:

1. The coordinate chart $\vec{\alpha}$ used to cover C_σ may degenerate when some $\alpha_i \to 0$ —such a degeneration would be signaled by the appearance of additional singularities in the Jacobian arising from the δ-functions in (6.1).
2. No *single* coordinate chart $\vec{\alpha}$ covers *all* of the boundaries of C_σ.

We can illustrate both points by considering a simple example. Recall from equation (5.14) the BCFW-bridge coordinates generated for the graph labeled by {4,6,5,7,8,9}:



(6.6)

Because the BCFW coordinates $\vec{\alpha}$ correspond to edge-variables, sending any $\alpha_i \to 0$ will have the effect of deleting the corresponding edge from the graph. The first subtlety mentioned above is reflected in the fact that some edge-variables—here, $\{\alpha_1, \alpha_2, \alpha_3, \alpha_6\}$—are attached to *irremovable* edges; the second subtlety is reflected in the fact that three of the seven removable edges—colored light gray in the figure—are *not* dressed with edge-variables. Of course, if we introduce additional $GL(1)$-redundancies at each vertex as we did in section 4.5, every removable edge could be dressed by a variable whose vanishing would give the corresponding boundary; this would make all the boundaries accessible, but at the cost of introducing vast redundancy.

A surprising fact—not very difficult to prove—is that *all* the boundaries of any cell $C \in G_+(k,n)$ can be found at the zero-locus of single coordinates in at least one chart from an atlas composed only of those charts generated by the BCFW-bridge construction (see section 3.2) in all its n cyclic manifestations (taking each of the n labels as the cyclic 'starting point' for the decomposition). To be clear, this claim only applies for the *specific scheme* described in section 3.2 used to decompose a permutation into adjacent transpositions—no other scheme is known to have this remarkable property.

7

The invariant top-form and the positroid stratification

We have seen that, associated with any d-dimensional cell of the positive Grassmannian, there is a natural associated form. In any of our natural coordinate charts, this d-form is just the "$d$log" measure,

$$\frac{d\alpha_1}{\alpha_1} \wedge \cdots \wedge \frac{d\alpha_d}{\alpha_d}, \tag{7.1}$$

which is a special case of a more general cluster volume discussed in Chapter 16. This form makes it obvious that boundary configurations are associated with residues for some $\alpha_i = 0$. It is also clear that we can view all cells $C \in G_+(k,n)$ as iterated residues of the *top-form* Ω^{top} on a *generic* configuration $C \in G_+(k,n)$.

A natural question is whether this top-form Ω^{top} can be written directly in terms of the 'matrix coordinates' c_a^α of C. In terms of matrix coordinates $C \equiv c_a^\alpha$, the desired measure $G(k,n)$ would have the form

$$\Omega = \frac{d^{k\times n}C}{\text{vol}(GL(k))} \frac{1}{f(C)}, \tag{7.2}$$

where $f(C)$ must be a function of the *minors* of C, and must scale uniformly as $f(tC) = t^{k\times n} f(C)$. Moreover, because the top-cell $G_+(k,n)$ *always* has precisely n codimension-one boundaries—corresponding to any k consecutive columns becoming linearly dependent—it is clear that $f(C)$ must have *at least* the n cyclic minors as factors:

$$f(C) = (1\cdots k)\cdots(n\cdots k{-}1) f'(C). \tag{7.3}$$

Because the product of the cyclic minors scales as $f(C)$ must, $f'(C)$ must be scale-invariant: $f'(tC) = f'(C)$. And so, $f'(C)$ can at most involve ratios of minors. However, any non-consecutive minors appearing as factors in $f'(C)$ would generate new, unwanted singularities for the top-cell—poles corresponding to codimension-one boundaries not in the positroid stratification—and any

consecutive minors in $f'(C)$ would make a double-pole, spoiling the logarithmic singularities corresponding to one of the *necessary* boundary configurations. Therefore, we are forced to conclude that the only choice is to take $f'(C) \to 1$. This means that the *only* viable ansatz for a measure on $G(k,n)$ with the desired properties is

$$\Omega = \frac{d^{k\times n}C}{\text{vol}(GL(k))}\frac{1}{(1\cdots k)\cdots(n\cdots k{-}1)}. \tag{7.4}$$

This strikingly simple form was first encountered in connection with "leading singularities" in reference [15].

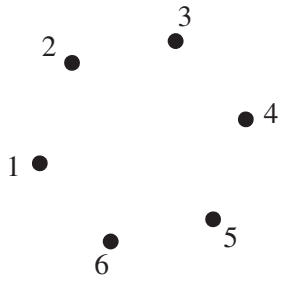
It is not hard to see the plausibility of a guess that $\Omega = \Omega^{\text{top}}$. We have just established that the poles of Ω and Ω^{top} are the same, and furthermore Ω does not have any zeroes on the Grassmannian. Thus $\Omega^{\text{top}}/\Omega$ is a function of the Grassmannian with no poles, and any such function must be a constant. So, we have

$$\frac{d^{k\times n}C}{\text{vol}(GL(k))}\frac{1}{(1\cdots k)\cdots(n\cdots k{-}1)} = \frac{d\alpha_1}{\alpha_1}\wedge\cdots\wedge\frac{d\alpha_{k(n-k)}}{\alpha_{k(n-k)}}. \tag{7.5}$$

This representation of the top-form will be crucial for most transparently seeing the dual conformal symmetry and Yangian invariance of the theory.

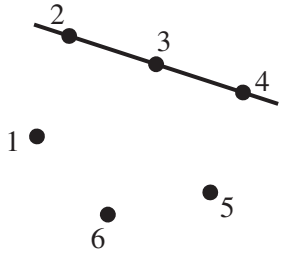
We will momentarily prove that $\Omega = \Omega^{\text{top}}$ by direct computation as well, but let us first step back and observe some remarkable properties of Ω. It is rather surprising that a form as simple as (7.4)—which has only n poles!—should be able to capture all of the intricate and beautiful structure of the positive Grassmannian in its iterated singularities. The reason why this isn't obviously impossible is that each of these n factors are generally kth-degree polynomials in the variables c_a^α, and whenever one such minor vanishes, other minors typically factorize, exposing further singularities and more structure below.

Let us consider an example that illustrates how the iterated factorizations of the consecutive minors expose all the cells in the positroid stratification. Consider the top-cell of $G(3,6)$,



$$\Leftrightarrow \frac{d^{3\times 6}C}{\text{vol}(GL(3))}\frac{1}{(123)(234)(345)(456)(561)(612)}. \tag{7.6}$$

Upon restricting this form to the residue where $(234) \to 0$, the configuration becomes:



(7.7)

Now, as described in Chapter 5, this configuration contains seven boundary configurations. How are we to see *seven* logarithmic singularities arising from the *five* remaining cyclic minors of (7.6)? The answer is simple: let us parameterize the pole $(234) \to 0$ by sending $c_3 \to \beta_2^3 c_2 + \beta_4^3 c_4$, under which the minors (123) and (345) each factorize:

$$\frac{1}{(123)(234)(345)(456)(561)(612)} \xrightarrow[c_3=\beta_2^3 c_2+\beta_4^3 c_4]{(234)\to 0 \text{ via}} \frac{1}{\underbrace{\beta_4^3(124)}_{(123)}\underbrace{\beta_2^3(245)}_{(345)}(456)(561)(612)},$$

exposing all seven of the boundary configurations! To further illustrate this point, let us now take a residue of this measure about the configuration setting $(561) \to 0$, by setting $c_6 \mapsto \beta_5^6 c_5 + \beta_1^6 c_1$; as before, this leads to the factorization of minors (456) and (612), leaving us with

$$\frac{1}{\beta_4^3(124)\beta_2^3(245)(456)(561)(612)} \xrightarrow[c_6=\beta_5^6 c_5+\beta_1^6 c_1]{(561)\to 0 \text{ via}} \frac{1}{\beta_4^3(124)\beta_2^3(245)\underbrace{\beta_1^6(461)}_{(456)}\underbrace{\beta_5^6(512)}_{(612)}},$$

which shows that this configuration has *eight* further boundary configurations. Proceeding in this way we can reconstruct all the cells of $G_+(3,6)$.

7.1 Equivalence with the canonical positroid volume form

In section 5.4 we showed that we can construct canonical coordinates for the top-cell of $G_+(k,n)$ recursively by first introducing coordinates

$$C^{(1,n)} \equiv \begin{pmatrix} \beta_1^1 & \beta_2^1 & \cdots & \beta_{n-1}^1 & \beta_n^1 \end{pmatrix}, \tag{7.8}$$

for $G(1,n)$, and then building up coordinates for any $G(k,n)$ recursively via:

$$C^{(k,n)} \equiv \begin{pmatrix} \beta_1^k \widehat{c}_1^{(k,n)} & \cdots & \beta_{n-1}^k \widehat{c}_{n-1}^{(k,n)} & 0 \\ \beta_1^k & \cdots & \beta_{n-1}^k & \beta_n^k \end{pmatrix} \text{ with } \widehat{c}_a^{(k,n)} \equiv \sum_{j=(a+1)}^{n} c_j^{(k-1,n)}. \tag{7.9}$$

Recall that these coordinates match those obtained by the BCFW-bridge construction upon the following trivial relabeling:

$$\begin{array}{|ccccccccc|}\hline \alpha_d & \alpha_{d-2} & \cdots & & \cdots & \cdots & \cdots & \alpha_\ell & \cdots\cdots \\ \alpha_{d-1} & \ddots & \ddots & & \ddots & \ddots & \alpha_{\ell+1} & \ddots & \ddots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & \cdots & \cdots & \alpha_{\ell+k-1} & \cdots & \cdots & \cdots\cdots & \alpha_3 & \alpha_1 \\ \hline \end{array} \Leftrightarrow \begin{array}{|cccc|}\hline \beta^1_{k+1} & \beta^1_{k+2} & \cdots & \beta^1_{n-1} & \beta^1_n \\ \beta^2_{k+1} & \beta^2_{k+2} & \cdots & \beta^2_{n-1} & \beta^2_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta^k_{k+1} & \beta^k_{k+2} & \cdots & \beta^k_{n-1} & \beta^k_n \\ \hline \end{array}$$

and the gauge-choice of setting the first k column vectors to the identity matrix. The motivation for relabeling the coordinates in this way is that the BCFW-coordinates give rise a gauge-fixed parameterization of $C(\beta^\alpha_a)$ of the form

$$\begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ k \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & k \end{array} \\ \left(\begin{array}{cccc|ccccc} 1 & 0 & \cdots & 0 & (\beta^1_{k+1}\cdots\beta^k_{k+1})+\cdots & (\beta^1_{k+2}\cdots\beta^k_{k+2})+\cdots & \cdots & (\beta^1_n\cdots\beta^k_n)+\cdots \\ 0 & 1 & \ddots & \vdots & (\beta^2_{k+1}\cdots\beta^k_{k+1})+\cdots & (\beta^2_{k+2}\cdots\beta^k_{k+2})+\cdots & \cdots & (\beta^2_n\cdots\beta^k_n)+\cdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \beta^k_{k+1} & \beta^k_{k+2} & \cdots & \beta^k_n \end{array}\right) \end{array} \begin{array}{c} \\ \\ \\ \\ \end{array} \begin{array}{c} \\ k+1 \\ \\ \\ \end{array} \cdots \begin{array}{c} \\ \\ \\ \\ \end{array} \begin{array}{c} \\ n \\ \\ \\ \end{array}$$

(7.10)

Here, we have highlighted terms to emphasize the fact that $c^\alpha_a \propto \beta^\alpha_a(\beta^{\alpha+1}_a\cdots\beta^k_a)+\ldots$, and that only this factor contributes to the Jacobian in going from coordinates c^α_a to coordinates β^α_a. In particular, it is easy to see that the entire Jacobian from this change of variables is simply

$$J \equiv \left|\frac{dc^\alpha_a}{d\beta^\alpha_a}\right| = \prod_{\alpha,a}\left(\beta^\alpha_a\right)^{\alpha-1}. \tag{7.11}$$

Somewhat less obviously, the cyclic minors are all simply expressed in these coordinates: each is the product of all the *highlighted* β^α_a in the lower-right triangle of the corresponding submatrix of (7.10):

$$(\ell\cdots\ell+k{-}1) = \prod_{\alpha=1}^{k}\left(\prod_{a=1}^{\alpha}\beta^\alpha_{(k+\ell-a)}\right) \Leftrightarrow \begin{vmatrix} \beta^1_\ell & \cdots & \cdots & \beta^1_{\ell+k-1} \\ \vdots & \ddots & \beta^2_{\ell+k-2} & \beta^2_{\ell+k-1} \\ \vdots & \ddots & \vdots & \vdots \\ \beta^k_\ell & \cdots & \beta^k_{\ell+k-2} & \beta^k_{\ell+k-1} \end{vmatrix}, \tag{7.12}$$

where the product of β's only ranges over relevant columns: $k+1\leq(k+\ell-a)\leq n$. And so, the product of all the consecutive minors is simply

$$(1\cdots k)(2\cdots k+1)\cdots(n\cdots k{-}1) = \prod_{\alpha,a}\left(\beta^\alpha_a\right)^\alpha. \tag{7.13}$$

Therefore, combining the product of all the cyclic minors with the necessary Jacobian given in (7.11) we have:

$$\frac{d^{k\times n}c^\alpha_a}{\mathrm{vol}(GL(k))}\frac{1}{(1\cdots k)\cdots(n\cdots k{-}1)} = \left(\prod_{\alpha,a}d\beta^\alpha_a\right)\frac{J}{\prod_{\alpha,a}\left(\beta^\alpha_a\right)^\alpha} = \prod_{\alpha,a}\frac{d\beta^\alpha_a}{\beta^\alpha_a} \tag{7.14}$$

as desired.

Let us briefly consider one concrete example of this equivalence. Consider the top-cell of $G(3,6)$, where the BCFW-bridge construction gives the matrix-representative

$$C(\alpha)=\begin{pmatrix} 1 & 0 & 0 & \alpha_9\alpha_8\alpha_6 & \alpha_7\alpha_5\alpha_3+\alpha_3\alpha_9(\alpha_5+\alpha_8) & \alpha_4\alpha_2\alpha_1+\alpha_1(\alpha_7(\alpha_2+\alpha_5)+\alpha_9(\alpha_2+\alpha_5+\alpha_8)) \\ 0 & 1 & 0 & -\alpha_8\alpha_6 & -\alpha_5\alpha_3-\alpha_3\alpha_8 & -\alpha_2\alpha_1-\alpha_1(\alpha_5+\alpha_8) \\ 0 & 0 & 1 & \alpha_6 & \alpha_3 & \alpha_1 \end{pmatrix},$$

which, upon relabeling the variables according to

$$\boxed{\begin{matrix} \alpha_9 & \alpha_7 & \alpha_4 \\ \alpha_8 & \alpha_5 & \alpha_2 \\ \alpha_6 & \alpha_3 & \alpha_1 \end{matrix}} \Longrightarrow \boxed{\begin{matrix} \beta_4^1 & \beta_5^1 & \beta_6^1 \\ \beta_4^2 & \beta_5^2 & \beta_6^2 \\ \beta_4^3 & \beta_5^3 & \beta_6^3 \end{matrix}}, \tag{7.15}$$

becomes

$$C(\beta)=\begin{pmatrix} 1 & 0 & 0 & \beta_4^1\beta_4^2\beta_4^3 & \beta_5^1\beta_5^2\beta_5^3+\cdots & \beta_6^1\beta_6^2\beta_6^3+\ldots \\ 0 & 1 & 0 & -\beta_4^2\beta_4^3 & -\beta_5^2\beta_5^3-\cdots & -\beta_6^2\beta_6^3-\cdots \\ 0 & 0 & 1 & \beta_4^3 & \beta_5^3 & \beta_6^3 \end{pmatrix}. \tag{7.16}$$

It is easy to see that the cyclic minors are given by,

$$\begin{aligned} (123) &= 1 & (456) &= \beta_6^1\beta_5^2\beta_6^2\beta_4^3\beta_5^3\beta_6^3 \\ (234) &= \beta_4^1\beta_4^2\beta_4^3 & (561) &= \beta_6^2\beta_5^3\beta_6^3 \\ (345) &= \beta_5^1\beta_4^2\beta_5^2\beta_4^3\beta_5^3 & (612) &= \beta_6^3 \end{aligned} \tag{7.17}$$

so that their product gives

$$(123)\cdots(612)=\left(\beta_4^1\beta_5^1\beta_6^1\right)^1\left(\beta_4^2\beta_5^2\beta_6^2\right)^2\left(\beta_4^3\beta_5^3\beta_6^3\right)^3; \tag{7.18}$$

and the Jacobian of going from c_a^α to β_a^α is easily seen to be

$$J\equiv\left|\frac{dc_a^\alpha}{d\beta_a^\alpha}\right|=\left(\beta_4^1\beta_5^1\beta_6^1\right)^0\left(\beta_4^2\beta_5^2\beta_6^2\right)^1\left(\beta_4^3\beta_5^3\beta_6^3\right)^2, \tag{7.19}$$

so that

$$\frac{d^{3\times 6}C}{\mathrm{vol}(GL(3))}\frac{1}{(123)(234)(345)(456)(561)(612)}=\prod_{\alpha,a}\frac{d\beta_a^\alpha}{\beta_a^\alpha}. \tag{7.20}$$

8

(Super-)conformal and dual conformal invariance

In this chapter, we will describe how the Grassmannian formulation of on-shell diagrams makes all the symmetries of the theory—both the super-conformal *and dual* super-conformal symmetries—completely manifest. Along the way, we will find it useful to recast the on-shell differential form's dependence on external kinematical data in a way that more transparently reflects the geometry of momentum-conservation; doing so, we will discover a correspondence between (some) cells $C \in G(k,n)$ with cells $\widehat{C} \in G(k-2,n)$.

8.1 The Grassmannian geometry of momentum conservation

Consider an arbitrary on-shell graph associated with the cell $\Gamma_\sigma \in G(k,n)$ labeled by the permutation σ associated with an on-shell differential form $f_\sigma^{(k)}(1,\ldots,n)$. Using any of the canonical coordinates for the cell $C(\alpha_1,\ldots,\alpha_d) \subset \Gamma_\sigma \in G(k,n)$, this form is given by:

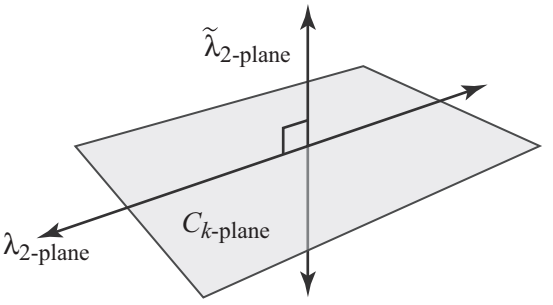
$$f_\sigma^{(k)} = \int \frac{d\alpha_1}{\alpha_1} \wedge \cdots \wedge \frac{d\alpha_d}{\alpha_d}\, \delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)\delta^{k\times 2}\big(C\cdot\widetilde{\lambda}\big)\delta^{2\times(n-k)}\big(\lambda\cdot C^\perp\big). \tag{8.1}$$

As we saw in Chapter 7, this can also be written as a residue of the top-form,

$$f_\sigma^{(k)} = \oint\limits_{C\subset\Gamma_\sigma} \frac{d^{k\times n}C}{\mathrm{vol}(GL(k))}\, \frac{\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)}{(1\cdots k)\cdots(n\cdots k-1)}\delta^{k\times 2}\big(C\cdot\widetilde{\lambda}\big)\delta^{2\times(n-k)}\big(\lambda\cdot C^\perp\big). \tag{8.2}$$

Recall from Chapter 4 that the (ordinary) δ-functions in (8.2) have the geometric interpretation of constraining the k-plane C to be *orthogonal to* the 2-plane $\widetilde{\lambda}$ and

to *contain* the 2-plane λ [15]:



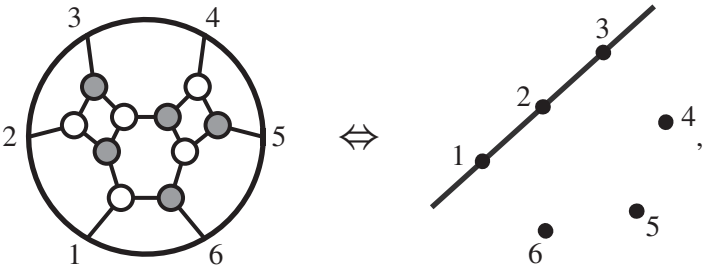
(8.3)

Because $\widetilde{\lambda} \subset \lambda^{\perp}$, 4 of the $2n(=2(n-k)+2k)$ constraints always represent momentum-conservation, leaving $(2n-4)$ constraints imposed on C in general. Therefore, cells of $G(k,n)$ with precisely $(2n-4)$ degrees of freedom can be *fully localized* by these constraints, and become *ordinary* super-functions of the external momenta; cells of lower dimension become functions with δ-function support, and cells of higher dimension represent *integration measures* on auxiliary, *internal* degrees of freedom (which may represent, for example, the degrees of freedom of internal loop-momenta).

The simplest example illustrating this localization is for $k=2$. Here the 2-plane C is just identified with the λ-plane, and equation (8.2) directly becomes the familiar Parke–Taylor formula for tree-level MHV super-amplitudes [70, 107]:

$$\begin{aligned}\mathcal{A}_n^{(2)} &= \int \frac{d^{2\times n}C}{\mathrm{vol}(GL(2))}\,\frac{\delta^{2\times 4}\big(C\cdot\widetilde{\eta}\big)}{(12)(23)\cdots(n1)}\delta^{2\times 2}\big(C\cdot\widetilde{\lambda}\big)\delta^{2\times(n-2)}\big(\lambda\cdot C^{\perp}\big)\\ &= \frac{\delta^{2\times 4}\big(\lambda\cdot\widetilde{\eta}\big)}{\langle 1\,2\rangle\langle 2\,3\rangle\cdots\langle n\,1\rangle}\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big).\end{aligned} \tag{8.4}$$

Let us look at a less trivial example of how this localization works for $k>2$. One of the on-shell diagrams contributing to the 6-particle $k=3$ tree-amplitude is



(8.5)

(see chapter 17), which is labeled by the permutation $\{3,5,6,7,8,10\}$. It is easy to see that (a $GL(3)$-representative of) the point C^* in this positroid cell that satisfies

the kinematical constraints is:

$$C^* = \begin{pmatrix} \lambda_1^1 & \lambda_2^1 & \lambda_3^1 & \lambda_4^1 & \lambda_5^1 & \lambda_6^1 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \lambda_4^2 & \lambda_5^2 & \lambda_6^2 \\ 0 & 0 & 0 & [56] & [64] & [45] \end{pmatrix}, \tag{8.6}$$

where $[a\,b] \equiv \det\{\widetilde{\lambda}_a, \widetilde{\lambda}_b\}$ is a minor of the matrix $\widetilde{\lambda}$. (Notice that $C^* \cdot \widetilde{\lambda} = 0$ because $\lambda \cdot \widetilde{\lambda} = 0$, and the third-row dotted-into $\widetilde{\lambda}$ gives an instance of (4.4).) Supported at this point, (8.2) generates the on-shell super-function, denoted $f^{(3)}_{\{3,5,6,7,8,10\}}$,

$$\frac{\delta^{3\times 4}(C^* \cdot \widetilde{\eta})\, \delta^{2\times 2}\big(\lambda \cdot \widetilde{\lambda}\big)}{\underbrace{\langle 23\rangle[56]}_{(234)|_{C^*}} \underbrace{(\langle 34\rangle[64] + \langle 53\rangle[56])}_{(345)|_{C^*}} \underbrace{s_{456}}_{(456)|_{C^*}} \underbrace{(\langle 61\rangle[64] + \langle 15\rangle[45])}_{(561)|_{C^*}} \underbrace{\langle 12\rangle[45]}_{(612)|_{C^*}}}, \tag{8.7}$$

where

$$s_{456} \equiv (p_4 + p_5 + p_6)^2 = \langle 45\rangle[45] + \langle 46\rangle[46] + \langle 56\rangle[56].$$

The particular $GL(3)$-representative of C^* given in (8.6) was chosen so that the Jacobian from all the δ-functions is 1, making the residue of (8.2) about the pole $(123) = 0$ easy to read off from C^*. Let us briefly mention that (8.7) makes *super* momentum-conservation manifest: in addition to the obvious $\delta^{2\times 2}\big(\lambda \cdot \widetilde{\lambda}\big)$ in (8.7), the (fermionic) δ-functions $\delta^{3\times 4}\big(C^* \cdot \widetilde{\eta}\big)$ includes the factor $\delta^{2\times 4}\big(\lambda \cdot \widetilde{\eta}\big)$—the supersymmetric extension of ordinary momentum conservation.

8.2 Twistor space and the superconformality of on-shell forms

In order to see the conformal symmetry of any theory, it is often wise to use *twistor* variables [108–112]. Not surprisingly then, it is *twistor* space—not momentum-space—which gives us the simplest basis in which to describe scattering amplitudes conformally. Formally, we go to twistor space by assuming that $\lambda, \widetilde{\lambda}$ are independent, real variables, and then Fourier-transform with respect to either the λ or $\widetilde{\lambda}$ variables [71]. It is not hard to see how this Fourier transform makes the action of conformal transformations particularly transparent. Working with spinor-helicity variables, the generators of translations, $P_{\alpha\dot{\beta}}$, Lorentz transformations, $J_{\alpha\beta}$ and $\overline{J}_{\dot{\alpha}\dot{\beta}}$, dilatations D, and special conformal transformations, $K_{\alpha\dot{\beta}}$, all look very different:

$$P_{\alpha\dot{\beta}} = \lambda_\alpha \widetilde{\lambda}_{\dot{\beta}},\ J_{\alpha\beta} = \frac{i}{2}\left(\lambda_\alpha \frac{\partial}{\partial \lambda^\beta} + \lambda_\beta \frac{\partial}{\partial \lambda^\alpha}\right), \text{ and } K_{\alpha\dot{\beta}} = \frac{\partial^2}{\partial \lambda^\alpha \partial \widetilde{\lambda}^{\dot{\beta}}}. \tag{8.8}$$

($\overline{J}$ is defined analogously to J.) However, if we Fourier-transform with respect to each of the λ's, say, using $\int d^{2\times n}\lambda\, e^{i\lambda \cdot \widetilde{\mu}}$, denoting the $(2 \times n)$ matrix of conjugate

variables by $\widetilde{\mu}$, the generators (8.8) become

$$P_{\dot{\alpha}\dot{\beta}} = i\widetilde{\lambda}_{\dot{\alpha}}\frac{\partial}{\partial\widetilde{\mu}^{\dot{\beta}}},\ J_{\dot{\alpha}\dot{\beta}} = \frac{i}{2}\left(\widetilde{\mu}_{\dot{\alpha}}\frac{\partial}{\partial\widetilde{\mu}^{\dot{\beta}}} + \widetilde{\mu}_{\dot{\beta}}\frac{\partial}{\partial\widetilde{\mu}^{\dot{\alpha}}}\right), \text{ and } K_{\dot{\alpha}\dot{\beta}} = i\widetilde{\mu}_{\dot{\alpha}}\frac{\partial}{\partial\widetilde{\lambda}^{\dot{\beta}}}. \quad (8.9)$$

(see [71] for a detailed discussion).

These are easy to recognize as the generators of $SL(4)$-transformations on *twistor* variables, denoted w_a, which combine $\widetilde{\lambda}$ and $\widetilde{\mu}$ according to:

$$w_a \equiv \begin{pmatrix}\widetilde{\mu}_a\\ \widetilde{\lambda}_a\end{pmatrix}. \quad (8.10)$$

Very nicely, under the action of the little group, the $\widetilde{\mu}$'s transform oppositely to the λ's so that the twistors transform uniformly like the $\widetilde{\lambda}$'s: $w_a \sim t_a^{-1} w_a$. Thus, we should view each w_a projectively as a point in $\mathbb{P}^3$. Furthermore, we can combine these ordinary variables w_a with the anti-commuting $\widetilde{\eta}$'s to form *super*-twistors $\mathcal{W}_a$ [113],

$$\mathcal{W}_a \equiv \begin{pmatrix}w_a\\ \widetilde{\eta}_a\end{pmatrix}, \quad (8.11)$$

for which the generators of the super-conformal group are simply those of $SL(4|4)$ —acting in the obvious way as super-linear transformations on the $\mathcal{W}$'s.

Now, given any of our on-shell forms, the Fourier transform with respect to the λ variables is straightforward as the only dependence on λ is in the term $\delta^{2\times(n-k)}\big(\lambda\cdot C^{\perp}\big)$. It will be useful to rewrite this to more directly reflect its geometric origin: the requirement that the plane C *contains* λ. Recall from section 4.3 (see equation (4.40)) that this means that there should exist a linear combination of the k row vectors of C that *exactly* match λ. In other words, if we parameterize such a linear combination by a $(2\times k)$ matrix ρ, we should be able to find a ρ for which $\rho\cdot C = \lambda$. Rewritten in terms of this auxiliary matrix ρ, the constraint that C contains λ becomes

$$\delta^{2\times(n-k)}\big(\lambda\cdot C^{\perp}\big) = \int d^{2\times k}\rho\ \delta^{2\times n}\big(\rho\cdot C - \lambda\big), \quad (8.12)$$

which makes it trivial to Fourier-transform to twistor space:

$$\int d^{2\times n}\lambda\ e^{i\lambda\cdot\widetilde{\mu}}\int d^{2\times k}\rho\ \delta^{2\times n}\big(\rho\cdot C-\lambda\big) = \int d^{2\times k}\rho\ e^{i(\rho\cdot C)\cdot\widetilde{\mu}} = \delta^{k\times 2}\big(C\cdot\widetilde{\mu}\big). \quad (8.13)$$

Therefore, in twistor space the constraints $\delta^{k\times 2}\big(C\cdot\widetilde{\lambda}\big)$ and $\delta^{2\times(n-k)}\big(\lambda\cdot C^{\perp}\big)$ together with the fermionic $\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)$ combine into the extremely elegant

$$\delta^{k\times 4}\big(C\cdot\widetilde{\eta}\big)\delta^{k\times 2}\big(C\cdot\widetilde{\lambda}\big)\delta^{k\times 2}\big(C\cdot\widetilde{\mu}\big) \Rightarrow \delta^{4k|4k}\big(C\cdot\mathcal{W}\big), \quad (8.14)$$

which makes the $SL(4|4)$-invariance of on-shell forms completely manifest. And so, in twistor space, the general on-shell form, (8.2), is simply

$$f_\sigma^{(k)} = \oint\limits_{C\subset\Gamma_\sigma} \frac{d^{k\times n}C}{\mathrm{vol}(GL(k))}\,\frac{\delta^{4k|4k}\big(C\cdot\mathcal{W}\big)}{(1\cdots k)\cdots(n\cdots k-1)}. \tag{8.15}$$

Note that our brief passage to twistor space was done mostly for formal reasons: in order to make the super-conformal symmetry of on-shell forms manifest. One disadvantage of this formalism, however, is that—at first glance—it appears that the integral over $C\in\Gamma_\sigma$ could be localized by all $4k$ (ordinary) δ-function constraints, while we know that on-shell forms associated with non-vanishing functions for generic (momentum-conserving) kinematical data correspond to $(2n-4)$-dimensional cells $\Gamma_\sigma\in G(k,n)$. The mismatch is due to the fact that Fourier-transforming to twistor space does not produce functions that are non-vanishing for a *generic* set of twistors. Instead, we get distributions on twistor space, imposing constraints on the twistor variables. Indeed, only $(2n-4)$ of the $4k$ δ-functions in (8.15) can be used to localize the Grassmannian integral while the remaining impose constraints on the configuration of external twistors.

8.3 Momentum-twistors and dual super-conformal invariance

In this subsection, we will review the arguments presented in [17] in order to discover that on-shell forms are quite surprisingly *also* invariant under an additional super-conformal symmetry. This new symmetry, called *dual* super-conformal invariance, combines with ordinary super-conformal symmetry to generate an infinite-dimensional symmetry algebra of on-shell forms known as *the Yangian* [114–117]. (Dual super-conformal invariance was first noticed in multi-loop perturbative calculations [118], and then at strong coupling [119]; this led to a remarkable connection between null-polygonal Wilson loops and scattering amplitudes—see, e.g., [119–127].)

Let us start by reconsidering the condition that the plane C contains the plane λ. Because this constraint is ubiquitous for on-shell forms, it is natural to sharpen our focus to the $(k-2)\equiv\widehat{k}$-plane—denoted $\widehat{C}$—that is the *projection of* C onto the orthogonal complement of λ. To be a bit more precise, suppose we have an operator $Q:\mathbb{C}^n\to\mathbb{C}^n$ with $\ker(Q)=\lambda$ such that,

$$Q\cdot\lambda=0. \tag{8.16}$$

With such an operator, we may define $\widehat{C}\equiv C\cdot Q$ so that $\widehat{C}\cdot\lambda=0$ trivially.

Now, super momentum-conservation is of course the statement that the planes $\widetilde{\lambda}$ and $\widetilde{\eta}$ are both in $\lambda^\perp$—which is the image of Q. And so we may use Q to express

$\widetilde{\lambda}$ and $\widetilde{\eta}$ in terms of some new, *generic* variables μ and η according to:

$$\widetilde{\lambda} \equiv \mu \cdot Q \quad \text{and} \quad \widetilde{\eta} \equiv \eta \cdot Q. \tag{8.17}$$

Defined in this way, any *unconstrained* planes μ and η will *automatically* define super momentum-conserving planes $\widetilde{\lambda}$ and $\widetilde{\eta}$.

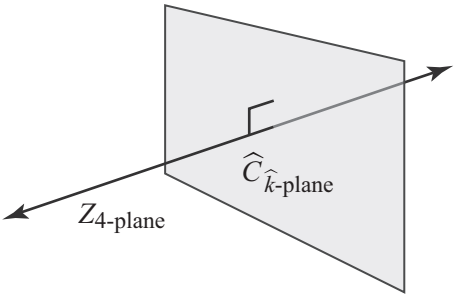
Let us now consider the constraint that C be orthogonal to the plane $\widetilde{\lambda}$. If Q were symmetric, then $C\cdot\widetilde{\lambda} = \widehat{C}\cdot\mu$; and similarly, $C\cdot\widetilde{\eta} = \widehat{C}\cdot\eta$. Putting all this together, the constraints imposed on the image $\widehat{k}$-plane $\widehat{C}$ would become simply

$$\delta^{\widehat{k}\times 2}\big(\widehat{C}\cdot\lambda\big)\delta^{\widehat{k}\times 2}\big(\widehat{C}\cdot\mu\big)\delta^{\widehat{k}\times 4}\big(\widehat{C}\cdot\eta\big) \Rightarrow \delta^{4\widehat{k}|4\widehat{k}}\big(\widehat{C}\cdot\mathcal{Z}\big), \tag{8.18}$$

where we have introduced the super *momentum-twistors* $\mathcal{Z}$ [128] according to:

$$\mathcal{Z}_a \equiv \begin{pmatrix} z_a \\ \eta_a \end{pmatrix} \quad \text{with} \quad z_a \equiv \begin{pmatrix} \lambda_a \\ \mu_a \end{pmatrix}. \tag{8.19}$$

Geometrically, the δ-functions $\delta^{k\times 4}\big(\widehat{C}\cdot Z\big)$ enforce that the plane $\widehat{C}$ be orthogonal to the 4-plane Z:



(8.20)

Notice that these δ-functions are invariant under a *new* $SL(4|4)$ symmetry, and thus it appears that we have uncovered a new super-conformal symmetry—one acting on the super-twistor variables $\mathcal{Z}_a$. However, there is one small catch: the measure of integration over the k-plane C does not necessarily descend to anything simple over the $\widehat{k}$-plane $\widehat{C}$. Indeed, depending on the choice of the projection operator Q, this resulting measure may have a complicated λ-dependence arising from the Jacobian of the change of variables from $(\widetilde{\lambda}, \widetilde{\eta})$ to (μ, η), and this dependence on λ may break the $SL(4)$ conformal symmetry.

But it turns out that for what is perhaps the most natural choice of a projection operator Q, everything works like magic. To better understand the scope of choices we could make in specifying Q, observe that such a projector can always be constructed via the Cramer's rule identities—the unique (up to rescaling) $(k+1)$-term identity satisfied by generic k-vectors. For a 2-plane λ, Cramer's rule

encodes the identities:

$$\lambda_a\langle bc\rangle + \lambda_b\langle ca\rangle + \lambda_c\langle ab\rangle = 0, \tag{8.21}$$

or equivalently, (if we prefer it to transform under the little group like $\widetilde{\lambda}_b$),

$$\lambda_a \frac{1}{\langle ab\rangle} + \lambda_b \frac{\langle ca\rangle}{\langle ab\rangle\langle bc\rangle} + \lambda_c \frac{1}{\langle bc\rangle} = 0. \tag{8.22}$$

If we combine any such n cyclically related identities, we will obtain a rank-$(n-2)$ matrix Q that projects onto $\lambda^\perp$. In order for Q to be *symmetric* as a matrix (which was necessary for $C\cdot\widetilde{\lambda}$ to be identified with $\widehat{C}\cdot\mu$), we must have λ_a and λ_c equally spaced about λ_b in (8.22). Of course, the most obvious and natural choice (and the only one that generates the magic we seek) would be to use the *consecutive* three-term identities:

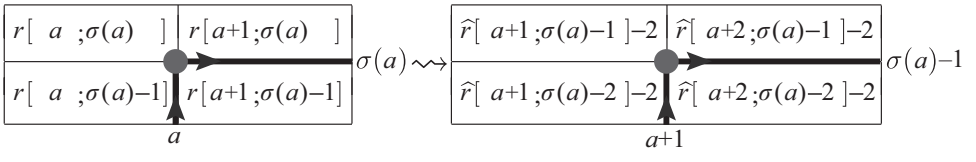
$$Q_{ab} \equiv \frac{\delta_{a-1\,b}\langle a\,a+1\rangle + \delta_{ab}\langle a+1\,a-1\rangle + \delta_{a+1\,b}\langle a-1\,a\rangle}{\langle a-1\,a\rangle\langle a\,a+1\rangle}. \tag{8.23}$$

For this choice of Q, it turns out that for any plane C containing λ, the plane $\widehat{C} \equiv C\cdot Q$ will have the property that for *any* consecutive chain of columns $\{c_a,\ldots,c_b\}$, $\text{span}\{\widehat{c}_a,\ldots,\widehat{c}_b\} \subset (\text{span}\{c_{a-1},\ldots,c_{b+1}\})$. That is, Q maps consecutive chains of columns onto consecutive chains of columns! An immediate consequence of this fact is that consecutive minors of C and $\widehat{C}$ are proportional to one another:

$$(1\,2\cdots k-1\,k)|_C = \langle 1\,2\rangle\langle 2\,3\rangle\cdots\langle k-1\,k\rangle\ (2\,3\cdots k-2\,k-1)|_{\widehat{C}}\,. \tag{8.24}$$

Thus, for this choice of Q—up to an overall λ-dependent factor (which combines with the Jacobian arising from changing variables $(\widetilde{\lambda},\widetilde{\eta})$ to (μ,η))—the top-form measure on $C\in G(k,n)$ given as the product of its consecutive minors, is mapped to the top-form on $\widehat{C}\in G(\widehat{k},n)$ of precisely the same form. And so, Q maps positroid cells in $G(k,n)$ (which contain a generic 2-plane λ) to positroid cells in $G(\widehat{k},n)$!

Conveniently, it turns out that the image of any cell $C\in G(k,n)$ in $G(\widehat{k},n)$ is very easy to identify by its permutation label. Because $\text{span}\{\widehat{c}_a,\ldots,\widehat{c}_b\} \subset (\text{span}\{c_{a-1},\ldots,c_{b+1}\})$, we have that $\widehat{r}[a;b] = r[a-1;b+1]-2$; and so, the entire table of ranks, (5.5), is preserved in going from C to $\widehat{C}$—merely shifted downward and to the right:



And so, a configuration $C_\sigma \in G(k,n)$ labeled by the permutation σ will be mapped to a configuration $\widehat{C}_{\widehat{\sigma}} \in G(\widehat{k},n)$ labeled by the permutation

$$\widehat{\sigma}(a) \equiv \sigma(a-1) - 1. \tag{8.25}$$

One last remarkable aspect of this change of variables is that the combination of all the λ-dependent factors arising from (8.24) when mapping the cyclic minors of $G(k,n)$ to cyclic minors of $G(\widehat{k},n)$ with the Jacobian of the change of variables from $(\widetilde{\lambda},\widetilde{\eta})$ to (μ,η) turns out to be none other than the Parke–Taylor (MHV) tree amplitude, (8.4)! And so,

$$f_\sigma^{(k)}(\lambda,\widetilde{\lambda},\widetilde{\eta}) = \frac{\delta^{2\times 4}\big(\lambda\cdot\widetilde{\eta}\big)\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big)}{\langle 1\,2\rangle\langle 2\,3\rangle\cdots\langle n\,1\rangle} \times f_{\widehat{\sigma}}^{(\widehat{k})}(\mathcal{Z}), \tag{8.26}$$

where

$$f_{\widehat{\sigma}}^{(\widehat{k})}(\mathcal{Z}) = \oint_{\widehat{C}\subset\Gamma_{\widehat{\sigma}}} \frac{d^{\widehat{k}\times n}\widehat{C}}{\mathrm{vol}(GL(\widehat{k}))} \frac{\delta^{4\widehat{k}|4\widehat{k}}\big(\widehat{C}\cdot\mathcal{Z}\big)}{(1\cdots\widehat{k})\cdots(n\cdots\widehat{k}-1)}. \tag{8.27}$$

This should not be too surprising, as the Parke–Taylor amplitude can be thought of as the most concise differential form consistent with super momentum conservation—and we know that any generic set of super-momentum-twistors $\mathcal{Z}$ give rise to data $(\widetilde{\lambda},\widetilde{\eta})$ that *manifestly* conserve super-momentum (this Grassmannian formula in terms of momentum twistors was introduced in [16]).

Let us briefly see how the dimensionality of cells $C_\sigma \in G(k,n)$ and their images $\widehat{C}_{\widehat{\sigma}} \in G(\widehat{k},n)$ are related. Because the rank of each chain $\widehat{r}[a+1;\widehat{\sigma}(a+1)]$ is lowered by 2 relative to $r[a;\sigma(a)]$, recalling the way dimensionality is encoded by the permutation (5.8), we see that

$$\begin{aligned} \dim(\widehat{C}_{\widehat{\sigma}}) &= \dim(C_\sigma) - 2n + k^2 - (k-2)^2 \\ &= \dim(C_\sigma) - (2n-4) + 4\widehat{k}; \\ \therefore\ \dim(\widehat{C}_{\widehat{\sigma}}) - 4\widehat{k} &= \dim(C_\sigma) - (2n-4). \end{aligned} \tag{8.28}$$

This is precisely as it should be: *generic* super momentum-twistors $\mathcal{Z}$ give rise to *generic* super-momentum conserving spinor-helicity data $\lambda,\widetilde{\lambda},\widetilde{\eta}$. Thus, the degree of the form $f_{\widehat{\sigma}}$ should be $\dim(\widehat{C}_{\widehat{\sigma}})$ minus the $4\widehat{k}$ ordinary δ-functions that force $\widehat{C}$ to be orthogonal to the generic 4-plane Z.

We should make one small point regarding the (existence of the) map between $G(k,n) \to G(\widehat{k},n)$: it is only well defined for cells C_σ *that contain a generic* 2*-plane* λ (this point is completely obvious from the geometry involved in the map's construction). In terms of the permutation σ which labels $C \in G(k,n)$, the criterion that C can contain a generic 2-plane λ translates into the statement that

$\sigma(a)-a\geq 2$ for all a. This guarantees that the permutation $\widehat{\sigma}$ is well-defined as an *affine* permutation, that is, that $\widehat{\sigma}(a)\geq a$. Suppose that instead we had $\sigma(a)=a+1$ for some a, then $c_a\in\text{span}\{c_{a+1}\}$, and so $\lambda\subset C$ would require that $\langle a\,a+1\rangle=0$. This all makes perfect sense, of course, because $\langle a\,a+1\rangle\to 0$ precisely corresponds to a singularity of the Parke–Taylor amplitude; and the Parke–Taylor amplitude being the *Jacobian* of the transformation to momentum-twistor space, any such singularity indicates that the change of variables is singular.

Let us conclude our discussion by illustrating the map to the 'momentum-twistor Grassmannian' for the example discussed above, (8.7), of the on-shell form associated with the cell in $G(3,6)$ labeled by the permutation $\{3,5,6,7,8,10\}$, (8.5). The image of this cell in the momentum-twistor Grassmannian $G(1,6)$ is labeled by $\widehat{\sigma}=\{3,2,4,5,6,7\}$. Since $\widehat{\sigma}(2)=2$, we have that $\widehat{c}_2=0$. A $GL(1)$-representative of the point $\widehat{C}^*$ that is orthogonal to the Z-plane in this cell is

$$\widehat{C}^*\equiv\big(\langle 3456\rangle\ 0\ \langle 4561\rangle\ \langle 5613\rangle\ \langle 6134\rangle\ \langle 1345\rangle\big), \tag{8.29}$$

where $\langle a\,b\,c\,d\rangle\equiv\det\{z_a,z_b,z_c,z_d\}$ is a minor of the matrix Z, and $\widehat{C}^*\cdot Z=0$ because of the 4-vector manifestation of Cramer's rule, (4.4). Supported on this point, (8.27) generates the momentum-twistor super-function,

$$f^{(1)}_{\{3,2,4,5,6,7\}}=\frac{\delta^{1\times 4}\big(\widehat{C}^*\cdot\eta\big)}{\underbrace{\langle 3456\rangle}_{(1)|_{\widehat{C}^*}}\underbrace{\langle 4561\rangle}_{(3)|_{\widehat{C}^*}}\underbrace{\langle 5613\rangle}_{(4)|_{\widehat{C}^*}}\underbrace{\langle 6134\rangle}_{(5)|_{\widehat{C}^*}}\underbrace{\langle 1345\rangle}_{(6)|_{\widehat{C}^*}}}. \tag{8.30}$$

And so, including the Parke–Taylor Jacobian, (8.26), we have:

$$f^{(3)}_{\{3,5,6,7,8,10\}}=\frac{\delta^{2\times 4}\big(\lambda\cdot\widetilde{\eta}\big)\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big)}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 45\rangle\langle 56\rangle\langle 61\rangle}\,\frac{\delta^{1\times 4}\big(\widehat{C}^*\cdot\eta\big)}{\langle 3456\rangle\langle 4561\rangle\langle 5613\rangle\langle 6134\rangle\langle 1345\rangle}.$$

9

Positive diffeomorphisms and Yangian invariance

We have seen that the map from twistor space to momentum-twistor space has a natural origin, providing an obvious geometric basis for dual conformal invariance. Let us now consider another obvious symmetry of the positive Grassmannian—namely, diffeomorphisms of Grassmannian coordinates that preserve the structure of the positroid stratification (equivalently, diffeomorphisms that leave measure on $G_+(k,n)$-invariant). Preserving the positive structure of the Grassmannian, we call this subset of diffeomorphisms *positive diffeomorphisms*. In this section, we illustrate the remarkable fact that the leading generators of infinitesimal positive diffeomorphisms directly match the level-one generators of the Yangian as described in [115] (see also [114, 116, 117, 129]).

Let us begin by broadly characterizing the infinitesimal diffeomorphisms in which we are interested. Consider any infinitesimal variation δC of $C \in G_+(k,n)$ that we may expand qualitatively as a power series,

$$\delta C \sim C + CC + CCC + \cdots . \tag{9.1}$$

We view a general infinitesimal diffeomorphism of C in terms of the variations δc_a^α for each matrix component of C. Because *positive* diffeomorphisms must preserve *all* positroid configurations, δc_a^α must vanish whenever c_a does; this restricts the class of diffeomorphisms to those of the form

$$\delta c_a^\alpha = \big(\Omega_a[C]\big)_\beta^\alpha c_a^\beta \quad \text{(no summation on } a\text{)}, \tag{9.2}$$

where each $\Omega_a[C]$ is itself expanded as a power series in the components of C. Considering $\Omega_a[C]$ as a $(k\times k)$ matrix, we may simplify our notation by writing:

$$\delta c_a = \big(\Omega_a[C]\big)\cdot c_a. \tag{9.3}$$

Note that any variation where Ω is proportional to the identity matrix is just an uninteresting (C-dependent) little group transformation. Note also that this variation takes the form of a different $GL(k)$ transformation on each column.

We can always use the global $GL(k)$-symmetry to bring the variation of any *one* column, say c_1, to zero: $\delta c_1 = 0$. (And without loss of generality, we can always take c_1 to be a non-vanishing column.)

Let us now determine what conditions must be imposed on $\Omega_a[C]$ in order to ensure that the variations δc_a preserve *all* positroid configurations. We will now demonstrate that there are no *nontrivial* variations in leading order in C, and that the first nontrivial positive diffeomorphisms—those quadratic in C—precisely correspond to the level-one generators of the Yangian as described in reference [115].

To leading order, each Ω_a is a C-independent $(k \times k)$ matrix. Consider any configuration for which $c_1 \propto c_2$, and let us use the $GL(k)$-symmetry to fix the variation of c_1 to zero. It is not hard to see that the only variation of c_2 that preserves the configuration in question would be the rescaling $\delta c_2 = t\, c_2$. This variation can be fully compensated by a little-group rescaling, allowing us to conclude that *no nontrivial* variation of c_2 is positive. Repeating this argument by starting with c_2 instead of c_1, and so on, we therefore see that the only *positive* leading-order diffeomorphisms are overall $GL(k)$-transformations and little group rescalings.

Nontrivial positive diffeomorphisms first arise at quadratic-order—when $\Omega_a[C]$ is linear in the components of C. Let us again consider any configuration for which $c_1 \propto c_2$, and use the $GL(k)$-symmetry to fix the variation of c_1 to zero. Because positive diffeomorphisms must preserve $r[a;b] \equiv \text{rank}\{c_a, \ldots, c_b\}$ generally—and $r[1;2]$ in particular—it is clear that the only allowed variations would be of the form,

$$\delta c_2 = (c_1 \omega_\beta^1) c_2^\beta \equiv c_1 (\omega^1 \cdot c_2). \tag{9.4}$$

We ignore any quadratic variation in c_2 as it represents a little-group rescaling. Here, ω_β^1 is an arbitrary k-vector parameterizing the variation. Notice that (9.4) is just a simple $GL(k)$-transformation of column c_2 by the matrix $M_\beta^\alpha \equiv (c_1^\alpha \omega_\beta^1)$. Applying the inverse of this transformation to *all* columns would of course trivialize $\delta c_2 \to 0$, allowing us to repeat the same logic to fix the most general form of δc_3, and so on. Continuing in this manner and then undoing each step's $GL(k)$-transformation so that we restore $\delta c_1 = 0$, the most general quadratic, positive diffeomorphism consistent with positivity would be of the form:

$$\begin{aligned}
&\delta c_1 = 0; \\
&\delta c_2 = c_1 (\omega^1 \cdot c_2); \\
&\delta c_3 = c_1 (\omega^1 \cdot c_3) + c_2 (\omega^2 \cdot c_3); \\
&\quad \vdots \\
&\delta c_n = c_1 (\omega^1 \cdot c_n) + \cdots + c_{n-1} (\omega^{n-1} \cdot c_n);
\end{aligned} \tag{9.5}$$

which we may summarize as:

$$\delta c_a = \sum_{b<a} c_b(\omega_\beta^b c_a^\beta). \tag{9.6}$$

We fixed the form of this transformation by demanding that the cells where $c_1 \propto c_2$ are left invariant, but quite nicely, we can see that this transformation leaves *all* cells invariant! Note that $r[1;a] \equiv \text{rank}\{c_1,\ldots,c_a\}$ is unchanged for all a, as the variations in (9.5) transform each c_a by factors proportional to columns that are always (trivially) spanned by the undeformed chains. And so, (9.6) preserves all $r[1;b]$—the entire first column of the table (5.5).

In order for the diffeomorphisms (9.6) to be *positive*, however, they must preserve the ranks $r[a;b]$ for *all* chains of columns; and so, we must find the subset of diffeomorphisms that are independent of our choice to single out δc_1. These can be found by continuing the sequence of successive variations in (9.5) back to δc_1, and requiring that this be consistent with our choice to fix $\delta c_1 = 0$:

$$\delta c_1 = c_1(\omega^1 \cdot c_1) + \cdots + c_n(\omega^n \cdot c_1) = \Big(\sum_{b=1}^n c_b \omega_\beta^b\Big) c_1^\beta = 0. \tag{9.7}$$

Because this must be satisfied for *all* configurations in $G_+(k,n)$, this must be independent of c_1^β. And so, the condition that ensures that (9.6) is positive is that,

$$\sum_{b=1}^n c_b \omega_\beta^b = 0. \tag{9.8}$$

This is simply the geometric statement that $\omega_\beta^a \subset C^\perp$ (for each index β separately). We have therefore constructed the most general set of infinitesimal, quadratic diffeomorphisms that preserve all cells in the positroid stratification of $G(k,n)$.

Recall that kinematical data—specified, say, in terms of super-twistor variables $\mathcal{W}$—are communicated to the Grassmannian via the constraint $\delta^{4k|4k}\big(C\cdot\mathcal{W}\big)$. This means that any symmetry-transformation acting on the $\mathcal{W}$'s can be recast as a transformation on the configuration C. In reference [115], it was shown that the level-one generators of the Yangian can be translated in this way to become symmetry generators acting on the matrix C by the operator

$$\mathcal{Q} \equiv \sum_{a=1}^n \mathcal{Q}_a \quad \text{with} \quad \mathcal{Q}_a \equiv \Big(\sum_{b<a} c_b^\alpha \mathcal{W}_I^b (\xi_\beta^I c_a^\beta)\Big) \frac{\partial}{\partial c_a^\alpha}, \tag{9.9}$$

which is easily seen to generate diffeomorphisms of the form

$$\delta c_a = \sum_{b<a} c_b w_I^b (\xi_\beta^I c_a^\beta), \tag{9.10}$$

which we immediately recognize as none other than the leading positive diffeomorphisms (9.6), where ω_β^b has been rewritten as

$$\omega_\beta^b \equiv w_I^b \xi_\beta^I, \tag{9.11}$$

for some (arbitrary) $(4 \times k)$-matrix ξ_β^I. Moreover, the condition on admissible variations, (9.8), is immediately seen to be *precisely* what is enforced by the constraint $\delta^{4k|4k}\big(C \cdot \mathcal{W}\big)$—which is imposed for all on-shell differential forms.

10

The kinematical support of physical on-shell forms

On-shell graphs with the right number of degrees of freedom to be completely localized for *generic*, (super-)momentum conserving kinematical data are obviously of particular interest. In momentum space, this requires that a configuration C associated with an on-shell graph admits solutions to both the constraint that it contains a generic two-plane $\lambda \in G(2,n)$, and is contained within the geometric-dual of another 2-plane $\widetilde{\lambda} \in G(2,n)$ satisfying $\lambda \cdot \widetilde{\lambda} = 0$. In terms of the permutation σ associated with an on-shell graph, these constraints minimally require that for any a, $(a+2) \leq \sigma(a) \leq (a+n-2)$. (Recall that the condition that $\sigma(a) \geq (a+2)$ is *necessary* for a configuration in $G(k+2,n)$ to even have a momentum-twistor dual in $G(k,n)$.) However, not all configurations that meet these conditions admit solutions to the combined constraints.

In this section, we will describe a purely combinatorial solution to the question of whether or not an on-shell graph vanishes for generic kinematical data; and if non-vanishing, how many solutions to the kinematical constraints exist. This turns out to be much simpler to do for the momentum-twistor Grassmannian rather than for configurations directly associated with on-shell graphs. This is partly because the kinematical constraints are much simpler for momentum-twistors than for the λ's and $\widetilde{\lambda}$'s.

Recall that when kinematical data is specified by momentum-twistors, $Z \in G(4,n)$, the configuration $\overline{C}_{\overline{\sigma}} \in G(k+2,n)$ *directly* associated with an $\mathrm{N}^k\mathrm{MHV}$ *on-shell graph* is mapped to its momentum-twistor image $\overline{C}_{\overline{\sigma}} \mapsto C_\sigma \in G(k,n)$, and the kinematical constraints become the simpler condition that $C \cdot Z = 0$. This imposes $4k$ constraints in general, and so we are most interested in $4k$-dimensional cells of $G(k,n)$, as these can be completely isolated by generic kinematical data. In terms of the orthogonal complement $Z^\perp$ of the twistors Z, the number of solutions to $C \cdot Z = 0$ is counted by the number of isolated points in $C \bigcap Z^\perp$.

As with any intersection-number problem in algebraic geometry, the solution can be found by decomposing both C and $Z^\perp$ into a homological basis for which the intersection numbers are known, such as Schubert cycles whose

intersection numbers are given by the Littlewood–Richardson rule (see [66]). The decomposition of (the closure of) an arbitrary positroid cell into Schubert cycles was recently presented in reference [40], and this provides us with a purely combinatorial answer to the 'number of intersections' question in which we are interested. And it turns out that for the special case of *generic* kinematical data, the machinery of [40] simplifies considerably. (We are thankful to Thomas Lam and David Speyer for helpful discussions regarding this specialization of the general case.)

A complete discussion of this story would require more space than warranted here; but let us briefly describe the ultimate, combinatorial solution to the question of kinematical support. The first step is to generalize our discussion slightly, and consider kinematical data specified for any number of dimensions:

Definition: For any $(m\times k)$-dimensional cell $C\in G_+(k,n)$, let $\Gamma^m(C)$ denote the number of isolated points in $C\bigcap Z^\perp$ for a *generic* m-plane $Z\in G(m,n)$.

The basic strategy is to define a *distinguished subset* $[\partial^k](C)$ of kth-degree boundary elements of C that contain non-overlapping subsets of the intersection points as projected to these boundaries, such that each element $C'\in[\partial^k](C)$ contains precisely $\Gamma^{m-1}(C')$ points. If this can be done, then $\Gamma^m(C)$ will be determined recursively by

$$\Gamma^m(C)=\sum_{C'\in[\partial^k](C)}\Gamma^{m-1}(C')\quad\text{with}\quad\Gamma^0(C)\equiv 1. \tag{10.1}$$

The magic, then, is entirely in the definition of the distinguished boundary elements $[\partial^k](C)$. Before we describe these in general, however, it may be helpful to build some intuition with two (very) simple cases for which (10.1) is easy to understand.

10.1 Kinematical support of NMHV Yangian-invariants

Although perhaps a bit trivial, it is worth noting that $\Gamma^m(C)=1$ for *all* m-dimensional configurations in $G(1,n)$—those relevant to NMHV amplitudes: given any generic m-plane Z, there is a unique configuration $C^*\in C\bigcap Z^\perp$ supplied by Cramer's rule, (4.4)—the unique $(m+1)$-term identity satisfied by generic m-vectors. This is of course obvious; but let us see what it suggests about how we may define the *distinguished* boundary elements $[\partial^1](C)$ that we seek to understand.

Just as $\Gamma^m(C)=1$ for any m-dimensional configuration in $G(1,n)$, $\Gamma^{m-1}(C)=1$ for any $(m-1)$-dimensional configuration. Therefore, in order for the recursive formula (10.1) to give us the right answer, we need only define $[\partial^1](C)$ to

systematically choose any *one* element of the boundary of C. One natural choice would be the configuration in ∂C that deletes the *first* non-vanishing column of C—that is, the boundary for which the maximum image of the configuration's permutation is 'raised.'

10.2 Kinematical support for one-dimensional kinematics

Let us now consider the slightly less trivial case of one-dimensional kinematics, where $Z \in G(1,n)$ and we are interested in finding $\Gamma^1(C)$ for k-dimensional configurations in $G(k,n)$. Unlike the situation for $k=1$, it is no longer the case that every k-dimensional configuration admits solutions to $C \cdot Z = 0$. The simplest example of a configuration for which $\Gamma^1(C) = 0$ occurs for the two-dimensional configuration in $G(2,4)$ labeled by the permutation $\{2,3,5,8\}$:

$$C(\alpha) = \begin{pmatrix} 1 & \alpha_1 & \alpha_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{10.2}$$

Notice that $C \cdot Z = 0$ implies that $z_4 = 0$, which is obviously not satisfied by a generic set of (one-dimensional) momentum-twistors. In contrast, consider the configuration labeled by the permutation $\{2,5,4,7\}$ represented by

$$C(\alpha) = \begin{pmatrix} 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & 1 & \alpha_2 \end{pmatrix}, \quad \text{for which} \quad C^* \equiv \begin{pmatrix} z_2 & -z_1 & 0 & 0 \\ 0 & 0 & z_4 & -z_3 \end{pmatrix} \tag{10.3}$$

is the unique solution to $C \cdot Z = 0$.

We can understand that a solution exists in the second case because each row of the matrix-representative of C has one degree of freedom—reducing each row to the simple case of $k=1$ described above. In the first example, however, no solution exists because the second row has no degrees of freedom—and so can itself be viewed as a zero-dimensional configuration in $G(1,n)$. Heuristically, then, in order for any solutions to $C \cdot Z = 0$ to exist, there must exist at least one degree of freedom in every row of any matrix-representative of C.

In terms of the permutation, the existence of a row without any degrees of freedom is indicated by any column a such that $\sigma(a) = a + n$. And so, a k-dimensional cell $C \in G(k,n)$ admits solutions to $C \cdot Z = 0$ for a generic 1-plane Z if and only if $\sigma(a) \neq a + n$ for all a.

10.3 A general combinatorial test of kinematical support

Combining the lessons learned from the two simple cases above, it is clear that solutions to $C \cdot Z = 0$ exist only if there are in some sense m degrees of freedom

in each row of any matrix-representative of C. A systematic way to test this combinatorially would be to find boundary elements of C that successively remove one degree of freedom from each row of C. Let us now describe how such boundary configurations can be found.

Recall that the lexicographically-first non-vanishing minor $A(\sigma)\equiv(a_1,\ldots,a_k)$ of any configuration C_σ is given simply by the images of σ that extend beyond n (see section 5.1). Because of this, we can always give a matrix-representative of C in the following, gauge-fixed form:

$$\begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ k \end{array}\!\!\begin{array}{c} \begin{array}{ccccccccccccc} \cdots & \cdots & a_1 & \cdots & \cdots & a_2 & \cdots & \cdots & a_k & \cdots & \cdots & \end{array} \\ \left(\begin{array}{ccccccccccccc} 0 & \cdots & 0 & 1 & * & \cdots & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 & * & \cdots & \cdots & \cdots & \cdots & * \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & * & \cdots & * \end{array}\right) \end{array}. \tag{10.4}$$

A boundary-element that removes one degree of freedom from the kth-row of C, for example, would be any that 'raises' a_k—that is, if $\sigma(c_k)=a_k$, then a boundary for which $\sigma'(c_k)=a'_k>a_k$. After this, a degree of freedom can be removed from the $(k-1)$th row, and so on. Notice, however, that at each stage, $A(\sigma)$ must be *raised*: if $A(\sigma)$ remained unchanged, then it would not indicate that a degree of freedom from any particular row had been removed, as we desire.

With this picture serving as motivation, we define the distinguished kth-degree boundary-elements of C, $[\partial^k](C)$, as follows. Let σ be the permutation labeling the configuration C, and let us define $A(\sigma)\equiv(a_1,\ldots,a_k)$—with $a_1<a_2<\cdots<a_k$—to be the images of σ that extend beyond n (the necklace $A^{(1)}(\sigma)$). Then $[\partial^k](C)$ is the set of all kth-degree boundaries of C obtained by a sequence of boundaries labeled by permutations $\sigma\xrightarrow{\partial}\sigma^{(1)}\xrightarrow{\partial}\sigma^{(2)}\xrightarrow{\partial}\cdots\xrightarrow{\partial}\sigma^{(k)}$ such that, *lexicographically*,

$$A(\sigma)<A(\sigma^{(1)})<A(\sigma^{(2)})<\cdots<A(\sigma^{(k)}). \tag{10.5}$$

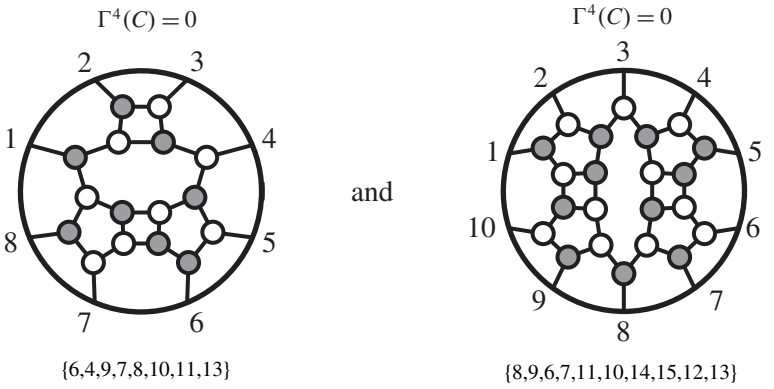
That is, if the configuration labeled by $\sigma^{(\ell-1)}$ with $A(\sigma^{(\ell-1)})=(a_1,\ldots,a_{k-\ell},\ldots,a_k)$ is found at the ℓth successive boundary, we take all $\sigma^{(\ell)}$ in its boundary for which $A(\sigma^{(\ell)})=(a_1,\ldots,a'_{k-\ell},\ldots,a_k)$ with $a_{k-\ell}<a'_{k-\ell}$.

In general, there can be many such elements of $[\partial^k](C)$, and each can contribute to $\Gamma^m(C)$. Putting all these contributions together, we find the recursive formula given above:

$$\Gamma^m(C)=\sum_{C'\in[\partial^k](C)}\Gamma^{m-1}(C')\qquad\text{with}\qquad\Gamma^0(C)\equiv 1. \tag{10.6}$$

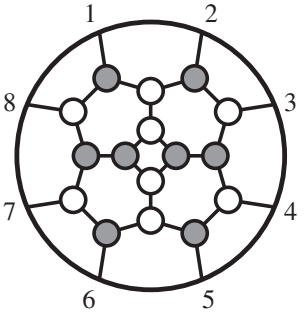
The utility of this combinatorial test is hard to overstate, as $4k$-dimensional cells in $G(k,n)$ with *non-vanishing* support become increasingly rare with large k and n. Cells with $\Gamma^4(C)=0$—for which $C\bigcap Z^\perp=\{\}$—represent generally vanishing functions that do not contribute to identities, for example. Many of

these fail the simple test of $(a+2) \leq \sigma(a) \leq (n+a+2)$, but with increasing frequency, configurations fail to have kinematical support for much more subtle reasons—demonstrating the value of having a more robust yet simple combinatorial test available. For example, neither of the following on-shell graphs—in $G(4,8)$ and $G(5,10)$, respectively—have kinematical support:



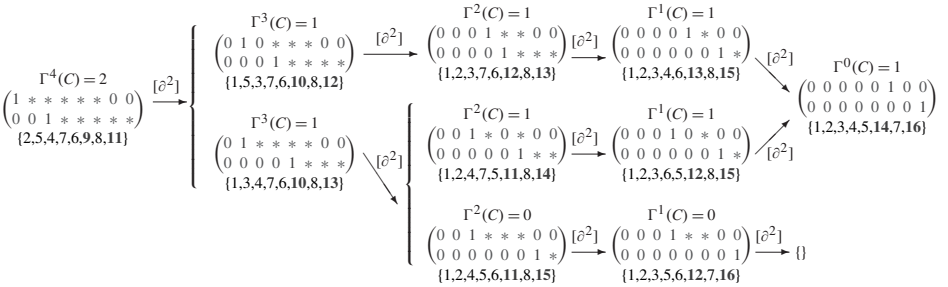
Configurations for which $\Gamma^4(C)=1$ correspond to manifestly *rational* functions of the kinematical data. More generally, however, when $\Gamma^4(C) > 1$ the isolation of internal degrees of freedom via $\delta^{k\times 4}\big(C\cdot Z\big)$ results in a (generally) *algebraic* function of the external twistors for each isolated solution $C^* \in C \bigcap Z^{\perp}$—each point giving us a Yangian-invariant that is individually of some physical interest. However, a highly nontrivial but general result is that the function obtained by summing over all isolated solutions to $C\cdot Z=0$ is *always rational*. Throughout the rest of this book, whenever we speak of 'the' function associated with a graph for which $\Gamma^4 > 1$—for example, when appearing in a identity (see Chapter 11)—we always implicitly mean the rational function obtained by summing over all particular solutions to $C\cdot Z=0$.

On-shell graphs that admit multiple solutions to the kinematical constraints are comparatively rare. The first on-shell graph for which more than one solution exists occurs for $G(4,8)$ and is well-known to physicists as the 'four-mass box':

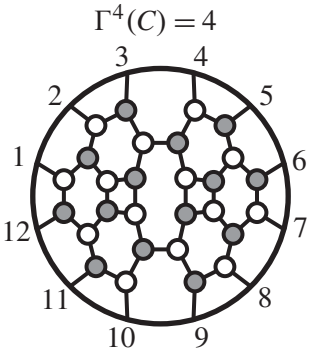


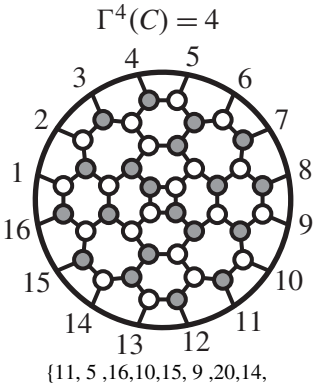
(10.7)

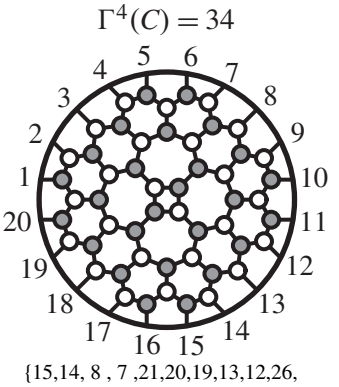
The image of this configuration in the momentum-twistor Grassmannian is labeled by $\{2,5,4,7,6,9,8,11\}$, for which we calculate $\Gamma^4(C)=2$ recursively as follows:



Configurations admitting more than two solutions to $C \cdot Z = 0$ are even rarer—and their rarity increases dramatically with increasing Γ^4. Indeed, almost no examples of Yangian-invariant functions for which $\Gamma^4(C) > 2$ were even known before the advent of the tools described in this section. But having the combinatorial test available allows us to systematically find and classify them. Three striking examples of on-shell graphs admitting many solutions to the kinematical constraints—for $G(6,12)$, $G(8,16)$, and $G(10,20)$, respectively—are:







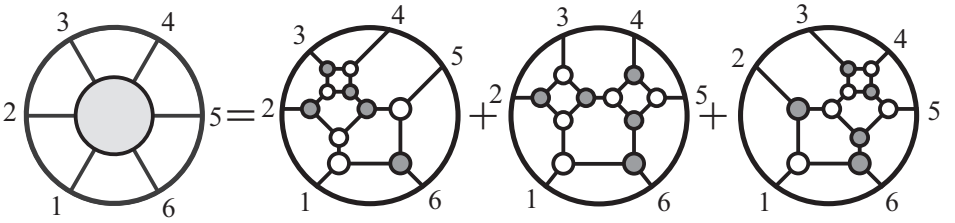
The on-shell diagram on the cover of this book, corresponding to a configuration in $C \in G(20,40)$, has $\Gamma^4(C) = 1024$ points of kinematical support.

11

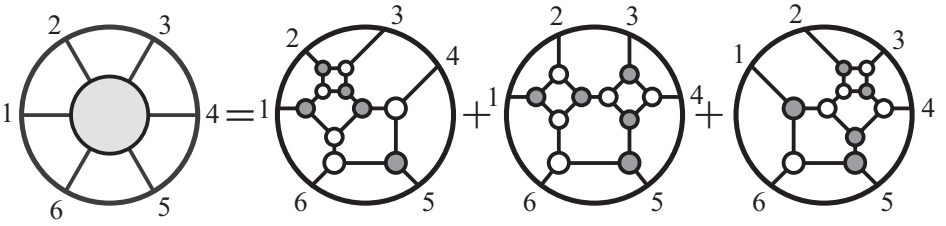
Homological identities among Yangian-invariants

In this section, we will focus primarily on on-shell differential forms for which the integral over auxiliary Grassmannian degrees of freedom is fully localized by the δ-function constraints, without imposing any conditions on the external kinematical data other than momentum conservation. These are on-shell diagrams with $(2n-4)$ degrees of freedom or their momentum-twistor images with $4k$ degrees of freedom, and for which $\Gamma^4(C) > 0$; we will refer to such on-shell forms as *Yangian-invariants*, and frequently refer to them (improperly) as 'functions' of the kinematical variables.

One of the most remarkable and important properties about Yangian-invariants is that they satisfy many intricate functional identities. Examples of such identities have long been known, and are crucial for our understanding of many important physical properties of scattering amplitudes. Perhaps the simplest and most familiar examples of such identities come from equating the various implementations of the BCFW recursion relations, (2.54); for example, for the six-particle NMHV tree-level scattering amplitude, BCFW recursion can alternatively lead to two distinct formulae depending on which pair of adjacent legs are singled out by the recursion: using the pair of legs (6 1), and we obtain the set of diagrams



while if we had chosen to recurse using the legs (56), we would have obtained



The identity generated by the equality of these two representations is not easy to prove directly if each term is viewed as a multivariate, rational 'function' of the kinematical data. However, its veracity is crucial to our understanding of many important properties of the complete amplitude. For example, although BCFW recursion breaks cyclicity by the choice of legs to deform, the entire amplitude—being cyclically invariant—must be independent of this choice.

A wide variety of such identities can be generated simply by equating all the myriad BCFW 'formulae' obtained by recursing the left- and right-amplitudes appearing across the BCFW-bridge in all possible ways (at each stage of the recursion). For example, for the eight-particle N^2MHV tree amplitude, there are many hundreds of ways to follow the recursion all the way down to a sum of 20 trivalent on-shell diagrams; this multitude of BCFW 'formulae' involves a total of 176 distinct Yangian-invariants in $G(4,8)$, and equating every pair of formulae leads to 74 linearly independent, 40-term identities satisfied among them.

Other than the equality of different BCFW formulae, however, few identities among Yangian-invariants were known until the Grassmannian formulation—the contour integral "$\mathcal{L}_{n,k}$"—was discovered [15]. But a complete understanding of the range of possible Yangian-invariants, and a systematic understanding of the relations they satisfy remained to be understood. In the remainder of this section, we will describe how *all* such identities arise *homologically* in the Grassmannian, and can be understood in purely geometric (even combinatorial) terms. In Chapter 13, we will illustrate the power of the combinatorial tools at our disposal, by giving an explicit classification of all Yangian-invariants and their relations through N^4MHV.

11.1 Homological identities in the Grassmannian

The six-term identity described above, which equates the two possible representations of the 6-particle $\mathrm{N}^{(k=1)}$MHV tree-amplitude, turns out to generate *all* the identities among NMHV Yangian-invariants. In order to see how this can be, let us first descend to the somewhat simpler situation that arises in the

momentum-twistor Grassmannian, where NMHV Yangian-invariants correspond to four-dimensional cells of $G(1,n)$.

All NMHV Yangian-invariants are essentially equivalent, as any four-dimensional configuration in $C \in G(1,n)$ involves precisely five non-vanishing 'columns'; and so, such configurations differ only in which of the five columns are involved. In terms of canonical coordinates, such a configuration would be represented by

$$C(\alpha) \equiv \begin{pmatrix} & & a & & b & & c & & d & & e & \\ \cdots & 0 & 1 & 0 \cdots 0 & \alpha_1 & 0 \cdots 0 & \alpha_2 & 0 \cdots 0 & \alpha_3 & 0 \cdots 0 & \alpha_4 & 0 \cdots \end{pmatrix},$$

and would be labeled by a permutation,

$$\sigma \equiv \begin{pmatrix} a & b & c & d & e \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ b & c & d & e & a \end{pmatrix}, \tag{11.1}$$

with $\sigma(j) = j$ for all other columns. Rather than labeling the configuration by its permutation, it is tempting to label it instead by its five non-vanishing columns—as a *5-bracket*, '$[abcde]$'. Given any generic momentum-twistors $Z \in G(4,n)$, there is a unique point $C^* \in C \bigcap Z^{\perp}$, which can be represented by the matrix

$$C^* \equiv \begin{pmatrix} & & a & & b & & c & & d & & e & \\ \cdots & 0 & \langle bcde\rangle & 0 \cdots 0 & \langle cdea\rangle & 0 \cdots 0 & \langle deab\rangle & 0 \cdots 0 & \langle eabc\rangle & 0 \cdots 0 & \langle abcd\rangle & 0 \cdots \end{pmatrix},$$

for which it is easy to see that $C^* \cdot Z = 0$ as an instance of Cramer's rule, (4.4). This leads to the general form of the essentially unique NMHV Yangian-invariant,

$$\underset{\{b,c,d,e,a\}}{[abcde]} \equiv \frac{\delta^{1\times 4}\big(\eta_a\langle bcde\rangle + \eta_b\langle cdea\rangle + \eta_c\langle deab\rangle + \eta_d\langle eabc\rangle + \eta_e\langle abcd\rangle\big)}{\langle bcde\rangle\langle cdea\rangle\langle deab\rangle\langle eabc\rangle\langle abcd\rangle}. \tag{11.2}$$

Notice that the 5-bracket $[abcde]$ as we have defined it is antisymmetric with respect to its arguments. This reflects the fact that the measure $d\log(\alpha_1)\wedge\cdots\wedge d\log(\alpha_4)$ is *oriented*. This 5-bracket is the simplest dual superconformal invariant and was first found in the literature in [130] in momentum space.

If we considered instead a five-dimensional configuration in $G(1,n)$, then the constraint $\delta^{1\times 4}\big(C\cdot Z\big)$ would fix only four of the internal degrees of freedom, leaving us with a one-dimensional integral over $G(1,n)$. In this case, Cauchy's theorem informs us that the sum of all the residues of this one-form will vanish. As each of these residues is itself a four-dimensional configuration of the form

above (11.2), this gives rise to an identity among five-brackets. Motivated by the notation used above, let us denote a generic 5-dimensional configuration in $G(1,n)$ by the *6-bracket* $[abcdef]$; then we find that

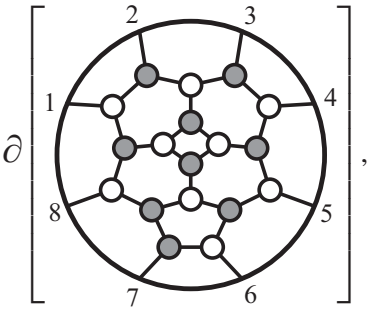
$$\underset{\{b,c,d,e,f,a\}}{\partial[abcdef]} = \underset{\{b,c,d,e,a,f\}}{[abcde]} - \underset{\{b,c,d,f,e,a\}}{[abcdf]} + \underset{\{b,c,e,d,f,a\}}{[abcef]} - \underset{\{b,d,c,e,f,a\}}{[abdef]} + \underset{\{c,b,d,e,f,a\}}{[acdef]} - \underset{\{a,c,d,e,f,b\}}{[bcdef]} = 0. \tag{11.3}$$

(Here, the signs are important: they reflect the fact that our formula for the 5-bracket (11.2) corresponds to a *particular orientation* of the four-dimensional cells; and so, when taking the boundary of $[abcdef]$ we must reorder the coordinates for each boundary cell accordingly—at the cost of introducing signs. Notice that the alternating signs here *precisely* capture the equality between two three-term, all-plus formulae as generated by equating BCFW formulae as described above.)

Notice that this six-term identity precisely reproduces the identity among 6-particle NMHV Yangian-invariants generated by equating BCFW recursion schemes. More importantly, however, because we understand that all NMHV Yangian-invariants are of the same basic form, the identity given above captures *all* the identities satisfied among NMHV Yangian-invariants.

The essential point in the example above is that if we consider a configuration $C \in G(k,n)$ *whose boundary* includes those associated with Yangian-invariant 'functions,' then the δ-function constraints will localize the Grassmannian integral to a one-dimensional integral, allowing us to use Cauchy's theorem to conclude that the sum of all the residues in the boundary will vanish; equivalently, that the combination of Yangian-invariants along any boundary $\partial(C)$ add to zero. This turns out to generate *all* the functional identities satisfied by Yangian-invariants, including many of impressive complexity.

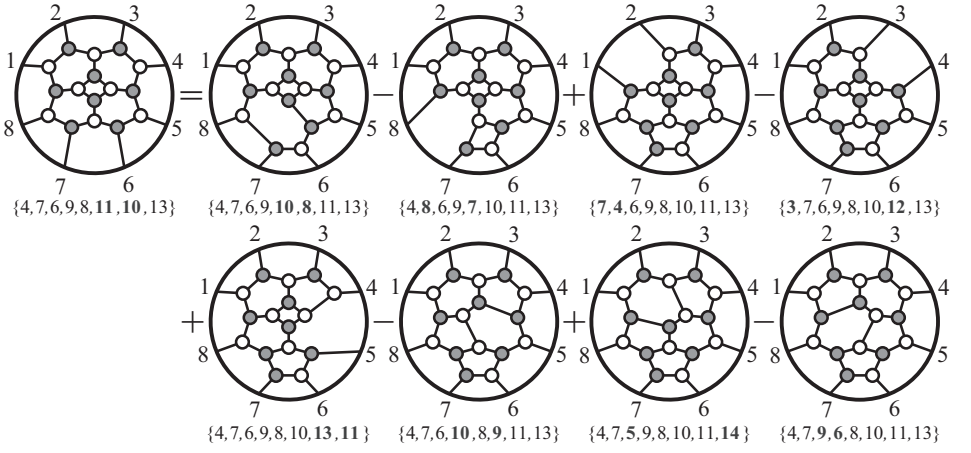
Recall that the boundary of an on-shell diagram is the collection of diagrams obtained by deleting its *removable* edges. And so we can find identities among N^{k}MHV on-shell differential forms by taking the boundary of any $(2n-3)$-dimensional cell in $G(k+2,n)$ for ordinary kinematical data, or any $(4k+1)$-dimensional cell of $G(k,n)$ for momentum-twistor kinematical data. One example of an identity found in this way generates an identity among 8-particle N^{2}MHV Yangian-invariants that is independent of all those identities found by equating various BCFW formulae, and can be understood as a way to represent the 'four-mass box' (which generally involves quadratic roots, as $\Gamma^4(C) = 2$) as a sum of purely rational Yangian-invariants: the boundary of the cell,



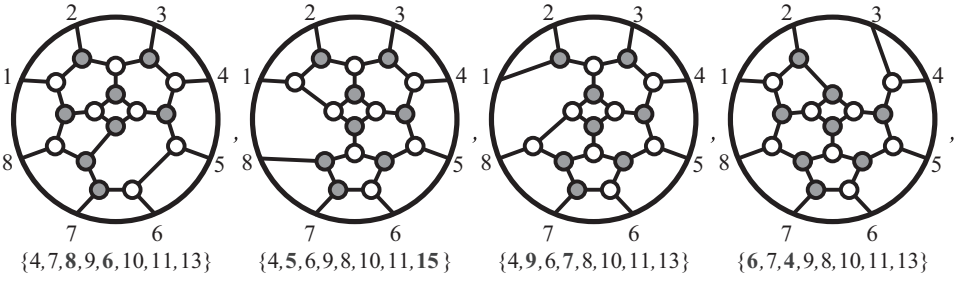
(11.4)

{4,7,6,9,8,10,11,13}

induces the following nine-term identity:



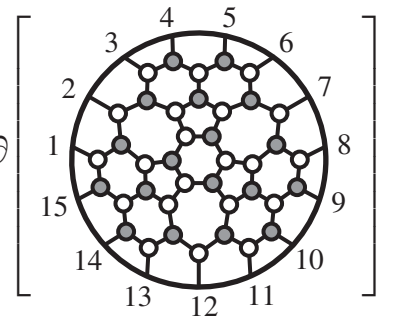
It is worth noting that we have only included the *non-vanishing* contributions to this identity—those graphs for which $\Gamma^4 > 0$; in addition to the nine graphs above, the boundary of {4,7,6,9,8,10,11,13} drawn in (11.4) also includes the graphs,



none of which have kinematical support ($\Gamma^4 = 0$), and so are generally vanishing functions of the kinematical data and therefore do not contribute to the identity.

Another particularly impressive example of an identity generated in this way is a 24-term identity among 15-particle N^{4}MHV Yangian-invariants, generated by

the boundary of the 27-dimensional cell,

$$\partial$$ 

(11.5)

{9,7,6,15,8,14,12,11,20,13,19,17,16,25,18}

the boundary of which includes eight cyclic classes of Yangian-invariants—three of which are quintic ($\Gamma^4(C) = 5$), two quartic ($\Gamma^4(C) = 4$), two quadratic ($\Gamma^4(C) = 2$), and one of which is rational ($\Gamma^4(C) = 1$).

12

(Relatively) orienting canonical coordinates charts on positroid configurations

Up to this point, we have largely taken for granted the fact that when we express an on-shell function $f_\sigma^{(k)}$ in terms of its Grassmannian representation,

$$f_\sigma^{(k)} = \int \frac{d\alpha_1}{\alpha_1} \wedge \cdots \wedge \frac{d\alpha_d}{\alpha_d}\, \delta^{k\times 4}\big(C_\sigma \cdot \widetilde{\eta}\big)\delta^{k\times 2}\big(C_\sigma \cdot \widetilde{\lambda}\big)\delta^{2\times(n-k)}\big(\lambda \cdot C_\sigma^\perp\big), \tag{12.1}$$

we should specify *which* canonical coordinates are used to parameterize $C_\sigma(\vec{\alpha})$—as different choices of coordinates may result in differently oriented volume forms, changing (12.1) by an overall sign. Because the sign of an on-shell function is not generally meaningful, there is often no 'preferred' choice of orientation for a given positroid's volume form. However, in order to combine (add) cells together—as in the identities discussed in Chapter 11, or when representing scattering amplitudes via recursion relations—we must ensure that the cells' volume forms are *relatively oriented* consistently.

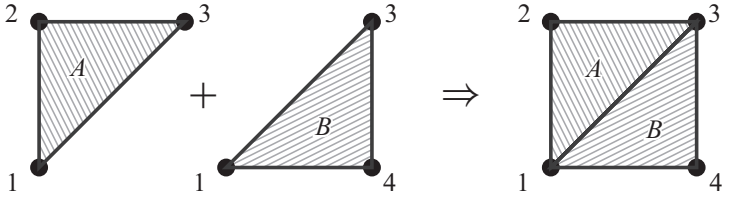
Before we describe how this may be done, we should note that not all combinations of cells can be consistently oriented. Nevertheless, it is not hard to see that the combinations of cells most interesting to us—those encoding identities among on-shell functions, and those representing scattering amplitudes—can always be consistently oriented. Moreover, *all* positroid cells generated by the BCFW recursion relations do have a 'preferred' orientation: the recursion relations directly endow each cell with an *oriented* volume—inherited from the orientations chosen for the three-point amplitudes. This will be discussed in detail below.

Whenever two positroid cells of the same dimension share a boundary, we say that they are *consistently oriented* if they induce oppositely-oriented volume forms on their shared boundary. The choice of orientation for a given cell *induces* orientations for all its boundaries in the following way. Suppose that we have chosen canonical coordinates for the configuration $C_\sigma(\vec{\alpha})$, and that the boundary configuration $C_{\sigma'} \in \partial(C_\sigma)$ is the zero-locus of some coordinate α_j; then the (oriented) volume form Ω_σ on $C_\sigma(\vec{\alpha})$ *induces* an oriented volume form on $C_{\sigma'}$

according to

$$\begin{aligned}\Omega_\sigma &\equiv \frac{d\alpha_1}{\alpha_1}\wedge\cdots\wedge\frac{d\alpha_{j-1}}{\alpha_{j-1}}\wedge\frac{d\alpha_j}{\alpha_j}\wedge\frac{d\alpha_{j+1}}{\alpha_{j+1}}\wedge\cdots\wedge\frac{d\alpha_d}{\alpha_d},\\ \mapsto \Omega_{\sigma'} &\equiv (-1)^j\frac{d\alpha_1}{\alpha_1}\wedge\cdots\wedge\frac{d\alpha_{j-1}}{\alpha_{j-1}}\wedge\frac{d\alpha_{j+1}}{\alpha_{j+1}}\wedge\cdots\wedge\frac{d\alpha_d}{\alpha_d}.\end{aligned} \tag{12.2}$$

Consider for example the pair of positroid configurations in $\{A,B\}\in G_+(1,4)$:



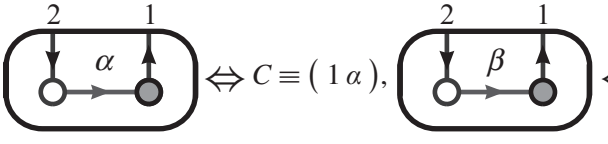
(12.3)

In order to combine the two regions and speak of the region "$A\cup B$," it must be that they induce oppositely oriented volume forms on their shared boundary. Indeed, the following coordinate charts on A and B are "consistently oriented" according to the definition above:

$$\begin{aligned}C_A(\vec{\alpha}) &\equiv \begin{pmatrix}1 & \alpha_2 & \alpha_3 & 0\end{pmatrix},\quad \Omega_A \equiv \frac{d\alpha_2}{\alpha_2}\wedge\frac{d\alpha_3}{\alpha_3}\xrightarrow[\alpha_2\to 0]{} -\frac{d\alpha_3}{\alpha_3}\equiv\Omega_{A\to(A\cap B)};\\ C_B(\vec{\alpha}) &\equiv \begin{pmatrix}1 & 0 & \alpha_3 & \alpha_4\end{pmatrix},\quad \Omega_B \equiv \frac{d\alpha_3}{\alpha_3}\wedge\frac{d\alpha_4}{\alpha_4}\xrightarrow[\alpha_4\to 0]{} +\frac{d\alpha_3}{\alpha_3}\equiv\Omega_{B\to(A\cap B)}.\end{aligned} \tag{12.4}$$

In the example above, the consistency of the orientations chosen for the configurations A and B was easy to check because their shared boundary was *accessible* from within each configuration's coordinate chart. Recall from the discussion in section 6.4 that a boundary configuration is said to be "accessible" within a given chart if it can be obtained as the zero-locus of a single coordinate. Although not all boundaries are accessible from any single chart, recall from section 6.4 that *every* boundary configuration is accessible from at least one chart in the *canonical atlas* consisting of the n lexicographical BCFW-bridge coordinate charts.

Unfortunately, however, not all the coordinate charts in the canonical atlas are equivalently oriented. To see this, consider the two charts in the canonical atlas for the top-cell of $G_+(1,2)$:



(12.5)

We can see that these two coordinate charts are oppositely oriented by noticing that a $GL(1)$-transformation of the second chart brings it to the form, $C(\beta)\simeq\begin{pmatrix}1 & \beta^{-1}\end{pmatrix}$,

which allows us to identify β with α^{-1}; and the Jacobian transforming $d\log(\alpha)$ into $d\log(\alpha^{-1})$ is (-1).

As illustrated by the example above, in order to find the orientations induced on all of the boundaries of a given positroid, we must first find a way to determine how the n charts in the canonical atlas are relatively oriented. We can actually do much better: in the following section, we will show how to determine the relative orientations of *any* pair of canonical coordinate charts for any positroid configuration. This will not only allow us to determine the relative orientations of the n charts in the canonical atlas and thereby allowing us determine how the orientation of one cell induces orientations on all its boundaries, but also to compare the orientations of the charts induced by two positroids' boundaries and thereby learn how to relatively orient two cells (which share a boundary) in order to combine them consistently.

12.1 Comparing the orientations of canonical coordinate charts

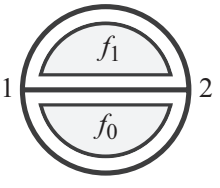
Without loss of generality, let us consider canonical coordinate charts corresponding to boundary measurements obtained from the face variables of a plabic graph as described in section 4.5. Although the BCFW-bridge coordinates are in fact edge-coordinates, because each bridge adds one new face there is a one-to-one correspondence (with unit Jacobian) between the BCFW-bridge variables and the face variables of the corresponding graph.

Recall from section 4.5 that a reduced plabic graph always has $n_F = \dim(\sigma) + 1$ faces. Thus, if we list the face variables $\vec{f} \equiv \{f_0, \ldots, f_d\}$, we can write the canonical volume form in terms of these variables as follows:

$$\begin{aligned} \Omega_\sigma(\vec{f}) &\equiv \Big(\frac{df_0}{f_0} \wedge \cdots \wedge \frac{df_d}{f_d}\Big)/\mathrm{vol}(GL(1)), \\ &\equiv (-1)^i \Big(\frac{df_0}{f_0} \wedge \cdots \wedge \frac{df_{i-1}}{f_{i-1}} \wedge \frac{df_{i+1}}{f_{i+1}} \wedge \cdots \wedge \frac{df_d}{f_d}\Big), \end{aligned} \tag{12.6}$$

with any choice of the redundant face f_i for elimination.

To see why this rule makes sense, consider the two coordinate charts for the top cell of $G_+(1,2)$ given above in (12.5). Each coordinate chart generates the same plabic graph, the faces of which we may choose to label as follows:



(12.7)

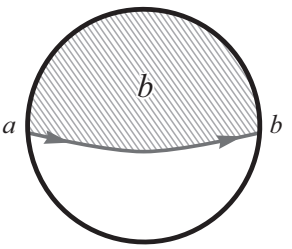
Notice that the first chart in (12.5) associates the coordinate α with f_1, while the second chart associates β with f_0. The rule above, (12.6), thus makes it obvious that the two charts are oppositely oriented.

Of course, as illustrated in the example above, merely reordering the face variables of a fixed plabic graph by an odd permutation will change the orientation of the resulting volume-form. More interestingly, as we saw in section 4.6, changing a graph by a square move also changes the orientation of the resulting volume-form (if the ordering of the face variables remain unchanged). And so, the relative orientation of two graphs' face-variable coordinates is simply determined by whether the number of square moves connecting them is odd or even.

But given two representative, reduced plabic graphs labeled by the same permutation, it can be quite hard to construct a sequence of moves that connects them (see, e.g., the pair of graphs drawn in equation (2.58)). Luckily, however, it turns out that we can *directly* compare the orientations of any pair of face coordinate charts. This is done by associating with each face variable a particular *tropical coordinate label*—the set of which will be the same for all representative graphs, making the relative orientation of the two charts be given simply by the signature of the relative ordering of their tropical coordinate labels.

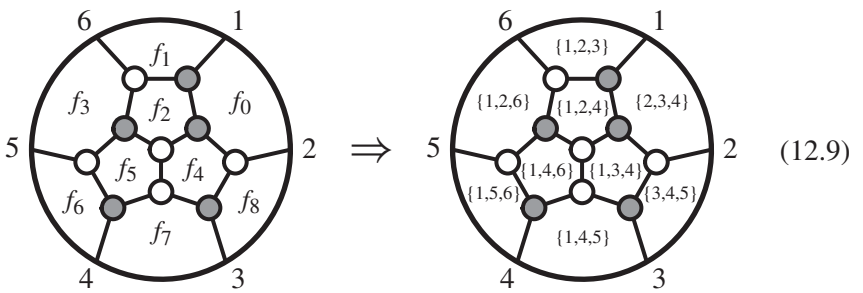
Constructing Tropical Labels for the Face Variables of a Plabic Graph

Let us describe how the faces of a plabic graph can be assigned such 'tropical' labels. The first step is to associate with each face f_i a *subset label*—an ordered k-tuple $\{f_i^\alpha\} \equiv \{f_i^1, \ldots, f_i^k\}$—defined in the following way. Let $\{f_i^\alpha\}$ denote the ordered set of all the images $b = \sigma(a)$ of left-right paths $\sigma : a \mapsto b$ whose counter-clockwise closure contains f_i—that is, $b \in \{f_i^\alpha\}$ for each f_i inside the shaded region below:



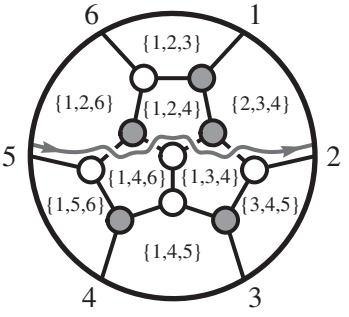
(12.8)

(It is a simple exercise to show that there are always k such paths for a graph corresponding to a configuration in $G_+(k,n)$.) More concretely, consider the following example of a graph associated with a positroid configuration in $G_+(3,6)$:



(12.9)

To see how the rule given above in (12.8) applies to the graph above, the following highlights all the faces whose subset labels contain 2:



(12.10)

Now, given the set of the subset labels for all the faces $\{f_0,\ldots,f_d\}$ of a given graph, we can associate a *tropical coordinate label* with each face as follows. For each face's subset label $\{f_i^1,\ldots,f_i^k\}$, there is a *maximal* element $\alpha\in\{1,\ldots,k\}$, such that for all f_j for which $f_j^\alpha = f_i^\alpha$, $f_j^\beta \geq f_i^\beta$ for all $\beta<\alpha$ and $f_j^\beta \leq f_i^\beta$ for all $\beta>\alpha$. We call the pair (α,f_i^α) the face f_i's *tropical label*. We can summarize this labeling as follows: starting from the faces $\{f_0,\ldots,f_d\}$ of a plabic graph, we construct each face's subset label, $f_i\Rightarrow\{f_i^1,\ldots,f_i^k\}$; and from the collection of all subset labels, we associate a tropical label (α,f_i^α) for each face:

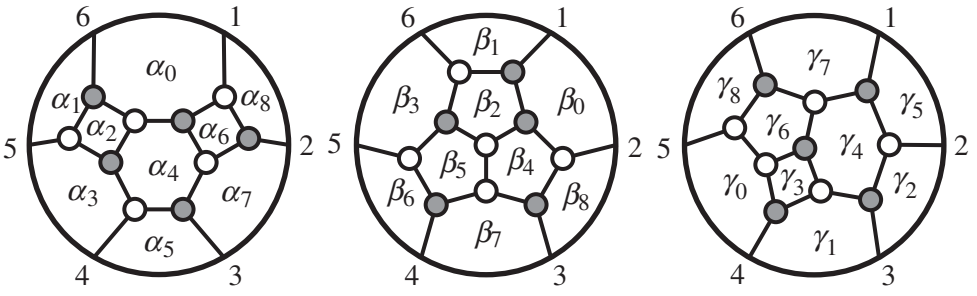
$$f_i\Rightarrow\{f_i^1,\ldots,f_i^\alpha,\ldots,f_i^k\}\Rightarrow(\alpha,f_i^\alpha). \tag{12.11}$$

For the example graph given in (12.9), the faces' tropical labels would be:

f_i	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
$\{f_i^\alpha\}$	{**2**,3,4}	{1,2,**3**}	{1,2,**4**}	{1,2,**6**}	{1,**3**,4}	{1,**4**,6}	{1,**5**,6}	{1,4,**5**}	{**3**,4,5}
(α,f_i^α)	(1,2)	(3,3)	(3,4)	(3,6)	(2,3)	(2,4)	(2,5)	(3,5)	(1,3)

The most important feature of these tropical labels (which we will not prove here) is that *the set of tropical labels for the faces of a plabic graph are the same for* all *representative graphs*. Thus, different sets of face coordinates are labeled by the same set of tropical labels, and the relative ordering between different sets of such tropical labels determines the relative orientation of the corresponding coordinates.

We can illustrate this more concretely with the following examples. Below are three representative plabic graphs labeled by the permutation $\sigma=\{4,5,7,6,8,9\}$:



(Notice that the middle graph above is the same as that of example (12.9).) Now, we can order the faces for each of the three graphs according to the ordering of their tropical labels as follows:

f_i	α_6	α_7	α_2	α_5	α_3	α_0	α_8	α_4	α_1
$\{f_i^\alpha\}$	{**2**,3,5}	{**3**,4,5}	{1,**3**,6}	{1,**4**,5}	{1,**5**,6}	{1,2,**3**}	{2,3,**4**}	{1,3,**5**}	{1,2,**6**}
(α, f_i^α)	(1,2)	(1,3)	(2,3)	(2,4)	(2,5)	(3,3)	(3,4)	(3,5)	(3,6)

f_i	β_0	β_8	β_4	β_5	β_6	β_1	β_2	β_7	β_3
$\{f_i^\alpha\}$	{**2**,3,4}	{**3**,4,5}	{1,**3**,4}	{1,**4**,6}	{1,**5**,6}	{1,2,**3**}	{1,2,**4**}	{1,4,**5**}	{1,2,**6**}
(α, f_i^α)	(1,2)	(1,3)	(2,3)	(2,4)	(2,5)	(3,3)	(3,4)	(3,5)	(3,6)

f_i	γ_5	γ_2	γ_6	γ_3	γ_0	γ_7	γ_4	γ_1	γ_8
$\{f_i^\alpha\}$	{**2**,3,4}	{**3**,4,5}	{1,**3**,6}	{1,**4**,6}	{1,**5**,6}	{1,2,**3**}	{1,3,**4**}	{1,4,**5**}	{1,2,**6**}
(α, f_i^α)	(1,2)	(1,3)	(2,3)	(2,4)	(2,5)	(3,3)	(3,4)	(3,5)	(3,6)

From this, it is easy to see (by the signatures of the relative orderings of the faces) that:

$$\Omega_\sigma(\vec{\alpha}) = +\Omega_\sigma(\vec{\beta}) = -\Omega_\sigma(\vec{\gamma}). \tag{12.12}$$

Before moving on, we should mention that adding a BCFW-bridge to a graph leaves the subset labels of all previous faces unchanged. Thus, BCFW-bridge coordinates can be labeled very nicely by subsets, making it very easy to compare the orientations of any BCFW-bridge coordinates for a given cell—in particular, the n coordinate charts in the canonical atlas.

12.2 Accessing boundary configurations from the canonical atlas

Now that we understand how the n lexicographical charts in the canonical atlas of BCFW-bridge coordinates are relatively oriented, it should be straightforward to determine the orientations of coordinates induced for all of a cell's boundary configurations. However, in order to do this, we must have a prescription for determining *which* boundary configurations are accessible from a given

coordinate chart in the atlas, and *which* configuration is at the zero-locus of *which* BCFW-bridge coordinate in the chart for which the boundary is accessible.

This turns out to be easier than it may at first appear. Recall that all the boundaries of a given configuration C_σ are labeled by permutations $\sigma' \equiv (a\,b)\circ\sigma$ related to σ by a transposition of the images of legs $(a\,b)$. As described in section 6.2, each such boundary corresponds to the removal of one 'crossing' in the juggling pattern. We can label these crossings by either the legs $\tau \equiv (a\,b)$ whose images are transposed, or by their images, $\widetilde{\tau} \equiv [\sigma(a)\,\sigma(b)]$. It will turn out to be very useful to label each boundary configuration of a given cell by the images $[\sigma(a)\,\sigma(b)]$ instead of the pre-images $(a\,b)$. The reason for this is that removing the lexicographically first BCFW-bridge from a graph will *always* leave these labels for all other boundaries unchanged. (This fact is not hard to prove; but we will leave it as an exercise for the reader.) Thus, although the crossings in the juggling pattern will move around as successive boundaries are taken, the labels of all other crossings will remain unchanged throughout the BCFW-bridge decomposition. In particular, this means that if we label each BCFW-bridge by the images being transposed, then the zero-locus of (any) bridge coordinate labeled by $[\sigma(a)\,\sigma(b)]$ will result in the boundary configuration labeled by $[\sigma(a)\,\sigma(b)]$.

To illustrate this, consider two of the charts in the canonical atlas for the permutation $\sigma = \{4,5,7,6,8,9\}$—the charts $C_\sigma(\beta)$ and $C_\sigma(\gamma)$ generated by the lexicographical BCFW-bridge decomposition starting with legs 2 and 5, respectively:

	τ	$\widetilde{\tau}$	2 3 4 5 6 7 ↓ ↓ ↓ ↓ ↓ ↓
β_8	(2 3)	[5 1]	5 7 6 8 9 10
β_7	(3 4)	[5 6]	7 5 6 8 9 10
β_6	(4 5)	[5 2]	7 6 5 8 9 10
β_5	(3 4)	[6 2]	7 6 8 5 9 10
β_4	(2 3)	[1 2]	7 8 6 5 9 10
β_3	(4 6)	[6 3]	8 7 6 5 9 10
β_2	(3 4)	[1 3]	8 7 9 5 6 10
β_1	(4 1)	[1 4]	8 9 7 5 6 10
			8 9 10 5 6 7

$$\begin{pmatrix} 0 & 1 & (\beta_4+\beta_8) & \beta_4(\beta_5+\beta_7) & \beta_4\beta_5\beta_6 & 0 \\ 0 & 0 & 1 & (\beta_2+\beta_5+\beta_7) & (\beta_2+\beta_5)\beta_6 & \beta_2\beta_3 \\ \beta_1 & 0 & 0 & 1 & \beta_6 & \beta_3 \end{pmatrix}$$

	τ	$\widetilde{\tau}$	5 6 7 8 9 10 ↓ ↓ ↓ ↓ ↓ ↓
γ_8	(5 6)	[2 3]	8 9 10 11 13 12
γ_7	(6 1)	[2 4]	9 8 10 11 13 12
γ_6	(5 6)	[3 4]	9 10 8 11 13 12
γ_5	(1 2)	[2 5]	10 9 8 11 13 12
γ_4	(6 1)	[3 5]	10 9 11 8 13 12
γ_3	(5 6)	[4 5]	10 11 9 8 13 12
γ_2	(1 3)	[3 1]	11 10 9 8 13 12
γ_1	(6 4)	[4 6]	11 10 13 8 9 12
			11 12 13 8 9 10

$$\begin{pmatrix} (\gamma_3\gamma_4+\gamma_3\gamma_7+\gamma_6\gamma_7) & \gamma_3\gamma_4\gamma_5 & 0 & 0 & 1 & (\gamma_3+\gamma_6+\gamma_8) \\ (\gamma_4+\gamma_7) & \gamma_4\gamma_5 & 0 & \gamma_1 & 0 & 1 \\ 1 & \gamma_5 & \gamma_2 & 0 & 0 & 0 \end{pmatrix}$$

(Notice that the representative plabic graphs generated by these two bridge decompositions correspond to the second two of the three examples given above—where each bridge coordinate has been identified with the corresponding face variable.) The configuration C_σ has a total of seven boundary configurations,

each of which is accessible in one of the following two charts:

$$\Omega_\sigma(\beta)\left\{\begin{array}{ll}\beta_8\to 0\text{:} & [5\,1]\equiv\{4,\mathbf{7},\mathbf{5},6,8,9\}\\ \beta_7\to 0\text{:} & [5\,6]\equiv\{4,\mathbf{6},7,\mathbf{5},8,9\}\\ \beta_5\to 0\text{:} & [6\,2]\equiv\{4,5,7,\mathbf{8},\mathbf{6},9\}\\ \beta_4\to 0\text{:} & [1\,2]\equiv\{4,5,\mathbf{8},6,\mathbf{7},9\}\end{array}\right\}=\partial(\{4,5,7,6,8,9\}) \tag{12.13}$$
$$\Omega_\sigma(\gamma)\left\{\begin{array}{ll}\gamma_8\to 0\text{:} & [2\,3]\equiv\{4,5,7,6,\mathbf{9},\mathbf{8}\}\\ \gamma_7\to 0\text{:} & [3\,4]\equiv\{\mathbf{3},5,7,6,8,\mathbf{10}\}\\ \gamma_3\to 0\text{:} & [4\,5]\equiv\{\mathbf{5},\mathbf{4},7,6,8,9\}\end{array}\right\}$$

As illustrated by the example above, it is often the case that only a few of the n charts in the canonical atlas are required to access all of a configuration's boundaries. Nevertheless, it is not hard to prove that any boundary configuration labeled by $[\sigma(a)\,\sigma(b)]$ will be accessible from the lexicographical BCFW-bridge coordinate chart where leg $\sigma(b)$ is taken as the minimum for the cyclic ordering. This guarantees every boundary is accessible from at least one chart in the canonical atlas, and allows us to constructively generate induced charts for all boundaries.

We now have everything required to systematically fix the relative orientations of combinations of cells required for identities or generated by the recursion relations. Because we know that all boundary configurations are accessible from the canonical atlas, and because the volume-form on any accessible boundary configuration is induced via equation (12.2), we can ensure that the relative orientation of any pair of cells that share a boundary are consistently chosen.

Notice that because every configuration in the codimension-two boundary of any configuration is reached as the boundary of precisely two codimension-one boundaries, this allows us to construct a *signed* boundary-operator ∂ which squares to zero symbolically (not merely modulo 2!): fix an arbitrary reference coordinate chart for every positroid strata, and say that $\partial:\sigma\mapsto\pm\sigma'$ if the reference, oriented chart on C_σ induces an orientation that agrees or disagrees with the reference chart on σ', respectively. The covering relation defined by ∂ is guaranteed to square to zero symbolically on the space of permutations.

13

Classification of Yangian-invariants and their relations

As we have seen, N^{k}MHV Yangian-invariant functions—expressed in terms of momentum-twistors—are determined by $4k$-dimensional configurations $C \in G(k,n)$ with non-vanishing kinematical support. As such, they are trivially classified by the permutations that label such cells. Importantly, because the dimensionality of these cells is independent of n, it turns out that for any fixed k, there *only* are only a *finite number of N^{k}MHV Yangian-invariant functions* with nontrivial dependence on the external momentum-twistors.

This is easy to see: consider any cell $C \in G(k,n)$ involving $\nu(C)$ non-vanishing columns; because $(\nu(C)-k)$ of the columns must each represent at least one degree of freedom, $\dim(C) \geq \nu(C)-k$; and so, for $4k$-dimensional cells, $\nu(C) \leq 5k$. Because any vanishing column c_a of C implies that $C \bigcap Z^{\perp}$ is geometrically-independent of the column z_a of Z, the Yangian-invariant associated with C will be functionally independent of the momentum-twistor z_a. And so, we immediately see that for any fixed k, the number of configurations $C \in G(k,n)$ with only nontrivial dependence on external momentum-twistors is strictly finite.

In addition to the finiteness of Yangian-invariant functions for fixed k, because *all* relations among them are generated as boundaries of $(4k+1)$-dimensional cells, their *relations* are also finite in number and can be similarly classified (for early work on classification of Yangian-invariants, see [15, 131–134]). Given the tools described in Chapter 10 to distinguish the permutations corresponding to cells with non-vanishing support, such a classification can be carried out relatively efficiently.

As described in section 11.1, N$^{(k=1)}$MHV Yangian-invariant functions correspond to cells $C \in G(1,n)$ with precisely five non-vanishing "columns"; because of this, there is only one NMHV Yangian-invariant function—up to trivial dependence on further twistors—which is commonly denoted by the *5-bracket* symbol $[1\,2\,3\,4\,5]$:

$$\frac{[1\,2\,3\,4\,5]}{\{2\,3\,4\,5\,6\}} \equiv \frac{\delta^{1\times 4}\big(\eta_1\langle 2345\rangle + \eta_2\langle 3451\rangle + \eta_3\langle 4512\rangle + \eta_4\langle 5123\rangle + \eta_5\langle 1234\rangle\big)}{\langle 2\,3\,4\,5\rangle\langle 3\,4\,5\,1\rangle\langle 4\,5\,1\,2\rangle\langle 5\,1\,2\,3\rangle\langle 1\,2\,3\,4\rangle}. \quad (13.1)$$

All other NMHV Yangian-invariants are equivalent to this object, but with the five non-vanishing columns being arbitrary, '$[abcde]$', (11.2). (A form very similar to this one was first introduced in [128] and used to develop a polytope description of NMHV tree-level amplitudes.) Similarly, there is a unique five-dimensional cell in $G(1,6)$ involving six columns, which we may denote $[1\,2\,3\,4\,5\,6]$. Its boundary generates the unique identity satisfied by 5-brackets (see section 11.1).

For N^{2}MHV, it is not hard to find that there are precisely 14 cyclically distinct Yangian-invariant functions (again, ignoring trivial dependence on further twistors). These are listed in Table 13.1, where we have also given canonical coordinates for each, and also expressed each as a product of 5-brackets—a notation somewhat more familiar to physicists, but which is generally cumbersome, as they frequently depend on geometrically defined momentum-twistors such as "$(12)\bigcap(345)$," which means "the momentum-twistor supported at the intersection of span$\{z_1,z_2\}$ with span$\{z_3,z_4,z_5\}$," found as a trivial application of Cramer's rule (4.4):

$$(12)\bigcap(345)\equiv z_1\langle 2345\rangle+z_2\langle 3451\rangle=-\big(z_3\langle 4512\rangle+z_4\langle 5123\rangle+z_5\langle 1234\rangle\big). \quad (13.2)$$

The configuration in Table 13.1 marked with a '$*$,' when lifted to an on-shell graph in $G(4,8)$, corresponds to the four-mass box (10.7). It is the unique N^{2}MHV function that admits more than one solution to the kinematical constraints. The two solutions to $C(\vec{\alpha})\cdot Z=0$ in this case are implicitly given in the 5-bracket form, due to the quadratic relation which defines the auxiliary twistors "A" and "B"; to be explicit, these auxiliary twistors are found as the solutions to

$$\left\{\begin{matrix}A\equiv z_7+\alpha z_8\\ B\equiv z_3+\beta z_4\end{matrix}\right\},\ \text{with}\ \alpha\equiv\frac{\langle 5637\rangle+\beta\langle 5647\rangle}{\langle 8563\rangle+\beta\langle 8564\rangle},\ \beta\equiv\frac{\langle 1273\rangle+\alpha\langle 1283\rangle}{\langle 4127\rangle+\alpha\langle 4128\rangle}. \quad (13.3)$$

Similarly, we may easily classify all the identities satisfied by N^{2}MHV Yangian-invariant functions, by classifying the nine-dimensional configurations of $G(2,n)$ whose boundaries generate such relations. It turns out that there are 24 cyclic classes of generators of identities—those listed in Table 13.2.

One interesting example of these is:

$$\begin{aligned}\partial\Big[\underset{\{2,5,4,7,6,8,9,11\}}{(1\,2)(3\,4)(5\,6)(7)(8)}\Big] \;=\;& \underset{\{2,5,4,7,6,\mathbf{9},\mathbf{8},11\}}{(1\,2)(3\,4)(5\,6)(7\,8)}-\underset{\{2,5,4,7,\mathbf{8},\mathbf{6},9,11\}}{(1\,2)(3\,4)(5)(7)(8)}+\underset{\{2,\mathbf{6},4,7,\mathbf{5},8,9,11\}}{(1\,2)(3\,4)(6)(7)(8)}\\ &-\underset{\{\mathbf{5},\mathbf{2},4,7,6,8,9,11\}}{(1)(3\,4)(5\,6)(7)(8)}+\underset{\{\mathbf{1},5,4,7,6,8,\mathbf{10},11\}}{(2)(3\,4)(5\,6)(7)\,(8)}-\underset{\{2,5,4,7,6,8,\mathbf{11},\mathbf{9}\}}{(8\,1\,2)(3\,4)(5\,6)(7)}\\ &+\underset{\{2,5,4,\mathbf{8},6,7,9,11\}}{(1\,2)(3\,4)(5\,6\,7)(8)}-\underset{\{2,5,\mathbf{3},7,6,8,9,\mathbf{12}\}}{(1\,2)(4)(5\,6)(7)(8)}+\underset{\{2,5,\mathbf{7},\mathbf{4},6,8,9,11\}}{(1\,2)(3)(5\,6)(7)(8)}\\ =\;&0.\end{aligned} \quad (13.4)$$

Table 13.1. *The complete classification of N^2MHV Yangian-invariant functions*

Projective Configuration of Column-Vectors	Canonical Coordinates for $C(\alpha)\in G_+(2,n)$	Written in Terms of 5-Brackets
$\underset{\{3,4,5,6,7,8\}}{(1)(2)(3)(4)(5)(6)}$	$\begin{pmatrix}1 & (\alpha_1+\alpha_2+\alpha_3+\alpha_4) & (\alpha_2+\alpha_3+\alpha_4)\alpha_5 & (\alpha_3+\alpha_4)\alpha_6 & \alpha_4\alpha_7 & 0\\ 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,(23)\cap(456),(234)\cap(56),6]$ $\times[2,3,4,5,6]$
$\underset{\{2,5,4,6,7,8,10\}}{(12)(34)(5)(6)(7)}$	$\begin{pmatrix}1 & \alpha_1 & (\alpha_2+\alpha_3+\alpha_4) & (\alpha_2+\alpha_3+\alpha_4)\alpha_5 & (\alpha_3+\alpha_4)\alpha_6 & \alpha_4\alpha_7 & 0\\ 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,(34)\cap(567),(345)\cap(67),7]$ $\times[3,4,5,6,7]$
$\underset{\{2,4,6,5,7,8,10\}}{(12)(3)(45)(6)(7)}$	$\begin{pmatrix}1 & \alpha_1 & (\alpha_2+\alpha_3+\alpha_4) & (\alpha_3+\alpha_4)\alpha_5 & (\alpha_3+\alpha_4)\alpha_6 & \alpha_4\alpha_7 & 0\\ 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,(345)\cap(67),7]$ $\times[3,4,5,6,7]$
$\underset{\{2,3,7,5,6,8,9,12\}}{(123)(456)(7)(8)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & (\alpha_3+\alpha_4) & (\alpha_3+\alpha_4)\alpha_5 & (\alpha_3+\alpha_4)\alpha_6 & \alpha_4\alpha_7 & 0\\ 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,(456)\cap(78),8]$ $\times[4,5,6,7,8]$
$\underset{\{2,3,5,8,6,7,9,12\}}{(123)(4)(567)(8)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & (\alpha_3+\alpha_4) & \alpha_4\alpha_5 & \alpha_4\alpha_6 & \alpha_4\alpha_7 & 0\\ 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,4,8]$ $\times[4,5,6,7,8]$
$\underset{\{2,3,6,5,8,7,9,12\}}{(123)(45)(67)(8)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & (\alpha_3+\alpha_4) & (\alpha_3+\alpha_4)\alpha_5 & \alpha_4\alpha_6 & \alpha_4\alpha_7 & 0\\ 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,(45)\cap(678),8]$ $\times[4,5,6,7,8]$
$\underset{\{2,3,6,5,7,9,8,12\}}{(123)(45)(6)(78)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & (\alpha_3+\alpha_4) & (\alpha_3+\alpha_4)\alpha_5 & \alpha_4\alpha_6 & 0 & 0\\ 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,(45)\cap(678),(456)\cap(78)]$ $\times[4,5,6,7,8]$
$\underset{\{2,3,5,7,6,9,8,12\}}{(123)(4)(56)(78)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & (\alpha_3+\alpha_4) & \alpha_4\alpha_5 & \alpha_4\alpha_6 & 0 & 0\\ 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,4,(456)\cap(78)]$ $\times[4,5,6,7,8]$
$*\underset{\{2,5,4,7,6,9,8,11\}}{(12)(34)(56)(78)}$	$\begin{pmatrix}1 & \alpha_1 & (\alpha_2+\alpha_3) & (\alpha_2+\alpha_3)\alpha_4 & \alpha_3\alpha_5 & \alpha_3\alpha_6 & 0 & 0\\ 0 & 0 & 1 & \alpha_4 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$\psi\,[A,1,2,3,4]$ $\times[B,5,6,7,8]$
$\underset{\{2,3,4,9,6,7,8,10,14\}}{(1234)(5678)(9)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_4\alpha_5 & \alpha_4\alpha_6 & \alpha_4\alpha_7 & 0\\ 0 & 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,4,9]$ $\times[5,6,7,8,9]$
$\underset{\{2,3,4,8,6,7,10,9,14\}}{(1234)(567)(89)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_4\alpha_5 & \alpha_4\alpha_6 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,4,(567)\cap(89)]$ $\times[5,6,7,8,9]$
$\underset{\{2,3,4,7,6,10,8,9,14\}}{(1234)(56)(789)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_4\alpha_5 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,4,(56)\cap(789)]$ $\times[5,6,7,8,9]$
$\underset{\{2,3,7,5,6,10,8,9,13\}}{(123)(456)(789)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_3\alpha_4 & \alpha_3\alpha_5 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & \alpha_4 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$\varphi\,[1,2,3,(45)\cap(789),(46)\cap(789)]$ $\times[(45)\cap(123),(46)\cap(123),7,8,9]$
$\underset{\{2,3,4,5,11,7,8,9,10,16\}}{(12345)(678910)}$	$\begin{pmatrix}1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8\end{pmatrix}$	$[1,2,3,4,5]$ $\times[6,7,8,9,10]$

$$\varphi\equiv\frac{\langle 45(123)\cap(789)\rangle\langle 46(123)\cap(789)\rangle}{\langle 1234\rangle\langle 4789\rangle\langle 56(123)\cap(789)\rangle},\ \psi\equiv\left(1-\frac{\langle A456\rangle\langle B812\rangle}{\langle A412\rangle\langle B856\rangle}\right)^{-1},\ \text{and}\ \begin{Bmatrix}A\equiv(78)\cap(56B)\\ B\equiv(34)\cap(12A)\end{Bmatrix}.$$

Here, we have only listed those boundary-configurations that have non-vanishing kinematical support. This identity lifts to the identity among on-shell graphs given in Chapter 11, which can be understood as expressing the (sum over particular solutions to $C \cdot Z = 0$ for the) four-mass box function as a sum of eight purely rational, N^{2}MHV Yangian-invariants.

Although it is not difficult to continue such a classification to higher and higher k, the space required to tabulate all objects soon grows overwhelming. Since it is of some interest, however, we summarize the statistical structure of the classification of Yangian-invariants and their identities through N^{4}MHV in Table 13.3.

This classification makes some interesting structures visible. For example, notice the extreme rarity of configurations with large Γ^4. Curiously we see that configurations with highest Γ^4 appear to be found when for $n = 3k + 2$ (this trend is known to continue well beyond $k = 4$).

Table 13.2. *A complete list of (cyclic generators) of identities among N^{2}MHV Yangian-invariants*

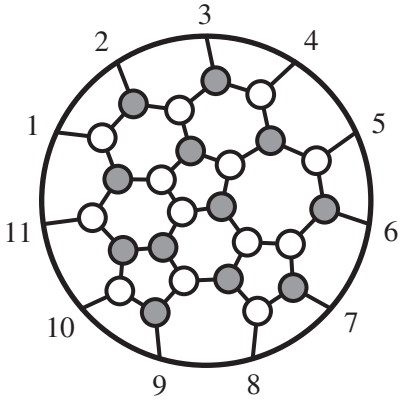
(12)(3)(4)(5)(6)(7) {2,4,5,6,7,8,10}	(123)(45)(6)(7)(8) {2,3,6,5,7,8,9,12}	(123)(4)(56)(7)(8) {2,3,5,7,6,8,9,12}
(123)(4)(5)(67)(8) {2,3,5,6,8,7,9,12}	(123)(4)(5)(6)(78) {2,3,5,6,7,9,8,12}	(12)(34)(56)(7)(8) {2,5,4,7,6,8,9,11}
(12)(34)(5)(67)(8) {2,5,4,6,8,7,9,11}	(1234)(567)(8)(9) {2,3,4,8,6,7,9,10,14}	(1234)(5)(678)(9) {2,3,4,6,9,7,8,10,14}
(1234)(5)(6)(789) {2,3,4,6,7,10,8,9,14}	(1234)(56)(78)(9) {2,3,4,7,6,9,8,10,14}	(1234)(56)(7)(89) {2,3,4,7,6,8,10,9,14}
(1234)(5)(67)(89) {2,3,4,6,8,7,10,9,14}	(123)(456)(78)(9) {2,3,7,5,6,9,8,10,13}	(123)(456)(7)(89) {2,3,7,5,6,8,10,9,13}
(123)(45)(678)(9) {2,3,6,5,9,7,8,10,13}	(123)(45)(67)(89) {2,3,6,5,8,7,10,9,13}	(12345)(6789)(10) {2,3,4,5,10,7,8,9,11,16}
(12345)(6)(7890) {2,3,4,5,7,11,8,9,10,16}	(12345)(678)(90) {2,3,4,5,9,7,8,11,10,16}	(12345)(67)(890) {2,3,4,5,8,7,11,9,10,16}
(1234)(5678)(90) {2,3,4,9,6,7,8,11,10,15}	(1234)(567)(890) {2,3,4,8,6,7,11,9,10,15}	(123456)(78910)(11) {2,3,4,5,6,12,8,9,10,11,18}

Table 13.3. *Numbers of cyclic classes of N^kMHV Yangian-invariant functions nontrivially involving n particles, and the number of cyclic classes of identities.*

		n																	
k		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	**Total**	**IDs**
1	total:	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1**
2	$\Gamma = 2$:	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
	$\Gamma = 1$:	0	1	2	5	4	1	0	0	0	0	0	0	0	0	0	0		
	total:	0	1	2	6	4	1	0	0	0	0	0	0	0	0	0	0	**14**	**24**
3	$\Gamma = 3$:	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0		
	$\Gamma = 2$:	0	0	0	0	2	15	27	19	2	0	0	0	0	0	0	0		
	$\Gamma = 1$:	0	0	1	6	54	177	298	274	134	30	1	0	0	0	0	0		
	total:	0	0	1	6	56	192	326	293	136	30	1	0	0	0	0	0	**1041**	**3669**
4	$\Gamma = 5$:	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0		
	$\Gamma = 4$:	0	0	0	0	0	0	0	6	40	60	30	3	0	0	0	0		
	$\Gamma = 3$:	0	0	0	0	0	0	1	47	179	232	125	12	0	0	0	0		
	$\Gamma = 2$:	0	0	0	0	0	10	177	1008	2630	3829	3158	1413	272	18	0	0		
	$\Gamma = 1$:	0	0	0	1	13	263	1988	7862	18532	28204	28377	18925	8034	2047	270	1		
	total:	0	0	0	1	13	273	2166	8923	21381	32334	31690	20353	8306	2065	270	1	**127776**	**603121**

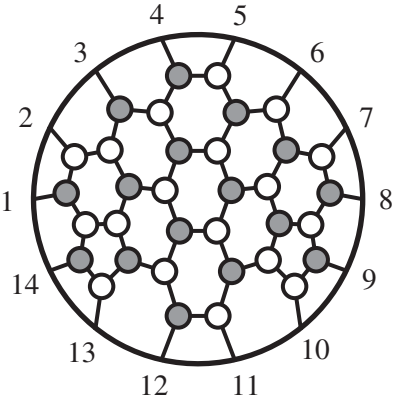
To illustrate the Yangian-invariants with exceptionally large Γ^4, below we give the on-shell graph that represents the unique 'cubic' N^3MHV Yangian-invariant, and one of those that represents a 'quintic' N^4MHV Yangian-invariant:

$\Gamma^4(C) = 3$



{4,9,6,11,8,14,10,13,12,18,16}

$\Gamma^4(C) = 5$



{6,5,13,12,11,9,10,18,17,15,14,22,16,21}

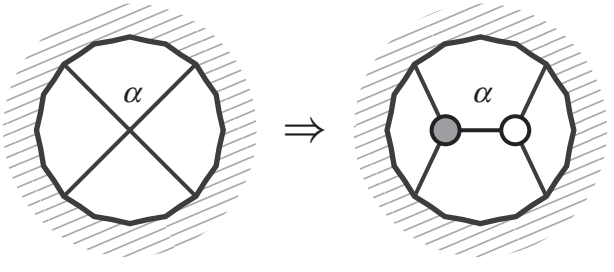
14

The Yang–Baxter relation and ABJM theories

We began our discussion linking permutations and scattering amplitudes in Chapter 3 by recalling the story of scattering in $(1+1)$-dimensional integrable theories (for a review, see [135]). In this chapter, we will see that this familiar story is actually contained as a special case of our new picture linking permutations to on-shell diagrams. And there is another special case that will turn out to give a theory of on-shell graphs for the ABJM theory [45] (see also [136–139]) defined in $(2+1)$ dimensions! Although both these stories are physically very rich on their own, we will content ourselves here by briefly sketching out the main points involved, leaving more detailed exposition and exploration to future work.

14.1 The on-shell avatar of the Yang–Baxter relation

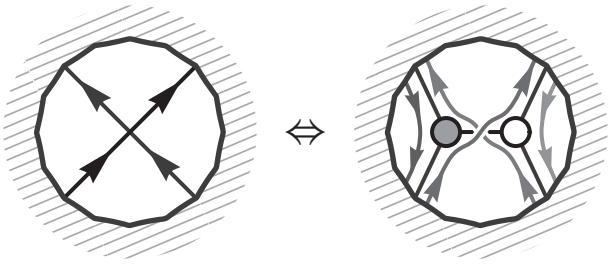
Recall the basic structure of the $(1+1)$-dimensional amplitudes from our discussion in Chapter 3, for which the fundamental interactions involved are 4-particle vertices. In order to relate these to our story, we must find a way to recast each 4-particle interaction (each carrying only one degree of freedom) in terms of an on-shell diagram with only trivalent vertices. The simplest way of doing this is to 'blow-up' each 4-point vertex according to the diagram



(14.1)

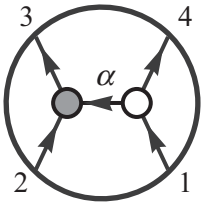
where only edges from blown-up vertices have weight different from unity.

Notice that the left-right path permutation moving from the bottom to the top of the graph agrees with the '$(1+1)$-permutation,' while the left-right path permutations from top to bottom are trivial:



(14.2)

Consider for example the 4-point vertex by itself,

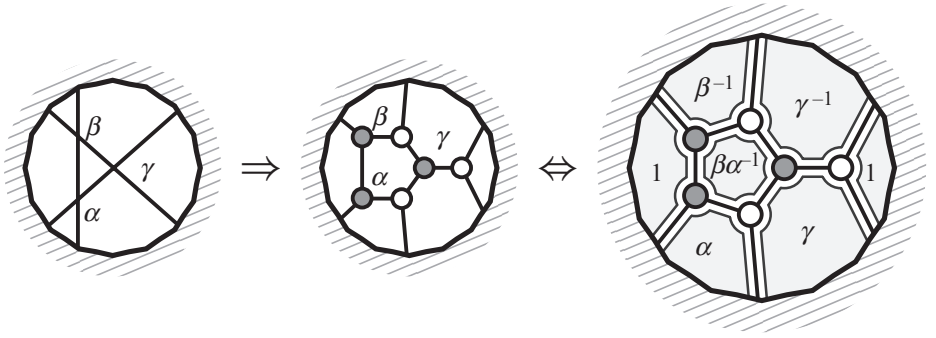


$$\Leftrightarrow\ C \equiv \left(\begin{array}{cc|cc} 1 & 0 & -\alpha & -1 \\ 0 & 1 & -1 & 0 \end{array}\right) \equiv (1_{2\times 2}|R_{12})\,, \qquad (14.3)$$

where we have given the point in the Grassmannian C obtained using edge variables and the perfect-orientation indicated in the figure (see section 4.5), from which we can read off the $2 \to 2$ "scattering matrix," which we have denoted R_{12}. In general, blowing up each 4-particle vertex allows us to translate any $(1+1)$-scattering diagram into a trivalent, on-shell diagram from which we can identify an $(n \times n)$ scattering matrix in the same way—identifying the point C in the Grassmannian in the gauge-fixed form

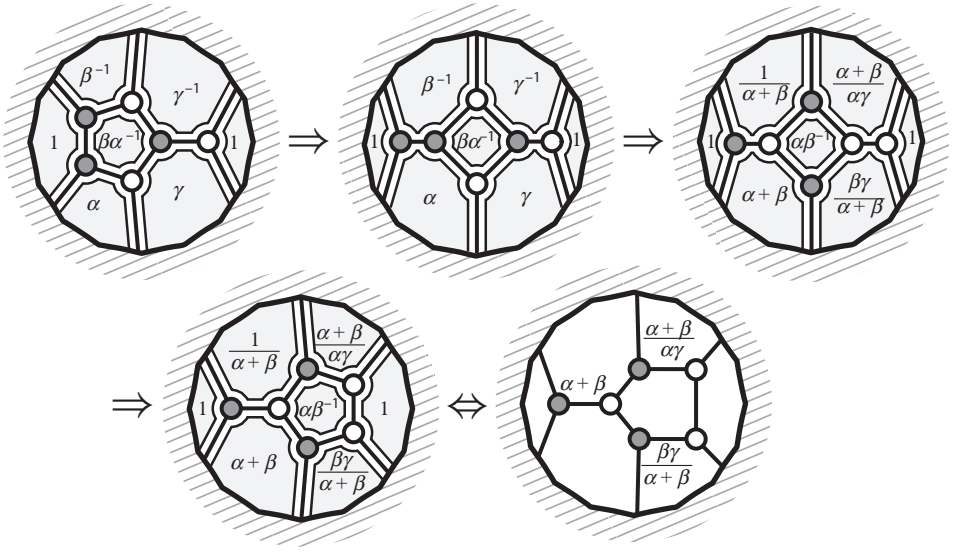
$$C_{n\times 2n} \equiv (1_{n\times n}|R_{n\times n}) \qquad (14.4)$$

As an example, let us look at the familiar configuration

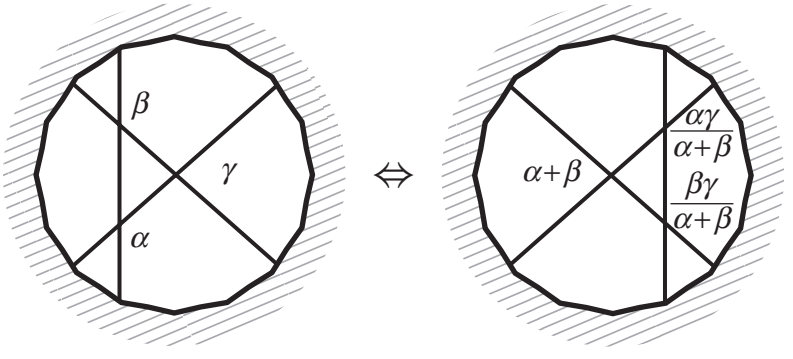


where on the right, we have recast the edge-variables into corresponding face-variables. Notice that because we are only putting nontrivial edge-weights on the "bridges" in the diagram, there are relations between the face variables.

Now, quite beautifully, we can see that the Yang–Baxter relation follows as a consequence of the more elementary actions of the merge-and-square moves! We can see this explicitly through the following sequence of moves, observing the effects induced on the face variables (see section 4.6):



From this, we may conclude that



This equivalence can be interpreted as a generalized Yang–Baxter relation for the R-matrices:

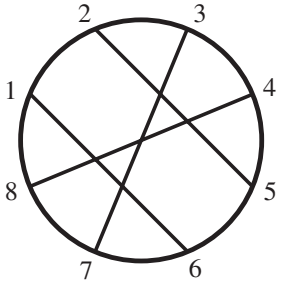
$$R_{12}(\beta)R_{13}(\gamma)R_{23}(\alpha) = R_{23}\left(\frac{\alpha\gamma}{\alpha+\beta}\right)R_{13}(\alpha+\beta)R_{12}\left(\frac{\beta\gamma}{\alpha+\beta}\right). \qquad (14.5)$$

In particular, if we set $\alpha+\beta=\gamma$, we recover the familiar Yang–Baxter equation:

$$R_{12}(\beta)R_{13}(\alpha+\beta)R_{23}(\alpha)=R_{23}(\alpha)R_{13}(\alpha+\beta)R_{12}(\beta). \tag{14.6}$$

14.2 ABJM theories

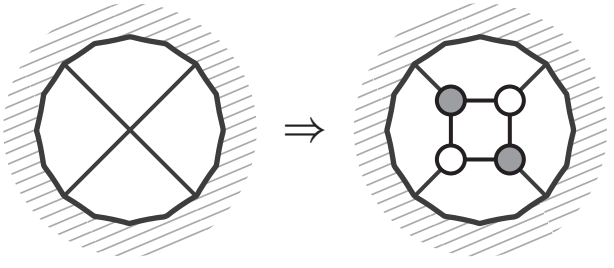
There is yet another natural way to associate a permutation with a scattering process. Suppose we have an even number, $2k$, of particle labels. We can divide them into two sets, A and B, of k elements each and draw arrows between them. Such a permutation takes some $a\to b$ and back via $b\to a$. We can then represent such a permutation graphically, with all labels on the boundary, as in the following:



A	B
1	6
2	5
3	7
4	8

(14.7)

We can then interpret this as an on-shell scattering process in a theory where each interaction is fundamentally a 4-particle vertex; and we can "blow up" each four-particle vertex into an element of $G(2,4)$, preserving the symmetrical nature of the permutation according to



(14.8)

This structure was also recently considered in [41]. As in the $(1+1)$-dimensional example, it is natural to try to associate each vertex with a single degree of freedom. Unlike the $(1+1)$-dimensional example, however, this restriction should keep us within the top-cell of $G(2,4)$. A very simple way of doing this would be to impose the restriction that the 2-plane is *null*; that is,

$$C\cdot C=0. \tag{14.9}$$

Notice that because the constraint $C\cdot C=0$ is *symmetric*, it represents $k(k+1)/2$ constraints in general; for $C\in G(2,4)$, this imposes only *three* constraints, leaving us with a single degree of freedom. In a canonical gauge, we can write:

$$C=\left(\begin{array}{cc|cc} 1 & 0 & is & -ic \\ 0 & 1 & ic & is \end{array}\right), \tag{14.10}$$

where $c\equiv\cos(\theta)$ and $s\equiv\sin(\theta)$ for some angle θ.

Exactly this Grassmannian structure has been found to represent scattering amplitudes for the $(2+1)$-dimensional ABJM theory [140–143]. As in $(3+1)$ dimensions, we can motivate the appearance of the Grassmannian by first looking at the geometry of external data. In $(2+1)$ dimensions, the momenta are grouped into a *symmetric* (2×2) matrix according to

$$p^{\alpha\beta}=\begin{pmatrix} p^0+p^2 & p^1 \\ p^1 & p^0-p^2 \end{pmatrix}, \tag{14.11}$$

so that null momenta are given by

$$p_a^{\alpha\beta}=\lambda_a^\alpha\lambda_a^\beta, \tag{14.12}$$

without any need for conjugate $\widetilde{\lambda}$'s. The Lorentz group acts as a single copy of $SL(2)$, so the λ_a are still represented by a 2-plane in n dimensions. However, momentum-conservation,

$$\sum\lambda_a^\alpha\lambda_a^\beta=0, \tag{14.13}$$

is now the statement that the λ-plane is orthogonal to itself. Thus, the external data is given not by a general point in $G(2,n)$, but by a point in the *null* Grassmannian of 2-planes in n dimensions. It is therefore not surprising to find the null Grassmannian playing a role in ABJM theory.

ABJM theories have $\mathcal{N}=6$ supersymmetries; if we diagonalize half of the supercharges, then the corresponding Grassmann coherent states are labeled by η^I for $I=1,\ldots,3$. Thus, the on-shell data can be collected into

$$\Lambda_a=\begin{pmatrix}\lambda_a\\ \eta_a\end{pmatrix}. \tag{14.14}$$

The ABJM amplitudes are not cyclically invariant, but *are* invariant under a cyclic shift by *two*. Notice that since we only have λ's, we lack the same little-group rescaling symmetry as we had in three dimensions; rather, we have only the symmetry of sending $\lambda_a\to-\lambda_a$, under which on-shell differential forms transform according to $f(-\Lambda_a)=(-1)^a f(\Lambda_a)$.

Let us now return to the basic 4-point vertex, and determine the natural measure on the space of null 2-planes in four dimensions. This space is easily seen to be

equivalent to $G(1,2) \simeq \mathbb{P}^1$: the two rows of a (2×4)-matrix can be viewed as four-vectors p_1, p_2 that are null and mutually orthogonal; we can therefore write,

$$p_1 = \lambda\,\widetilde{\lambda}_1, \qquad p_2 = \lambda\,\widetilde{\lambda}_2, \tag{14.15}$$

and use the $GL(2)$-freedom to write $\widetilde{\lambda}_1 \equiv (1\,0)$, $\widetilde{\lambda}_2 \equiv (0\,1)$, and $\lambda \equiv (1\,z)$. This demonstrates the equivalence of the null Grassmannian $C \subset G(2,4)$ with $\mathbb{P}^1$, and also provides us with a natural measure: $d\log(z)$. Using this identification, we can write the null-plane $C \subset G(2,4)$ in terms of z according to

$$\begin{pmatrix} -i & -iz & z & 1 \\ -z & 1 & -i & iz \end{pmatrix}. \tag{14.16}$$

Performing a $GL(2)$-transformation to recast this matrix-representative of C in a canonical gauge brings it to the form given above in (14.10), with the identification

$$s = \frac{2z}{z^2+1}, \quad \text{and} \quad c = \frac{z^2-1}{z^2+1}. \tag{14.17}$$

In terms of the natural measure $d\log(z)$ on the null subspace, the fundamental 4-point interaction in the ABJM theory can then be represented by

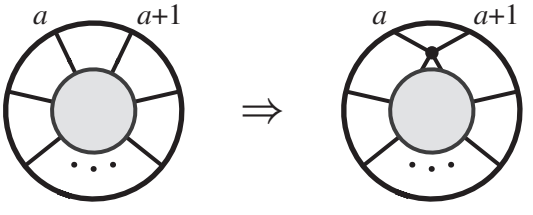
$$\mathcal{A}_4 = \int \frac{dz}{z}\, \delta^{4|6}(C(z)\cdot\Lambda); \tag{14.18}$$

equivalently, we may view this as having been obtained from a measure defined on all of $G(2,4)$, but restricted to the null subspace by the constraint $\delta^3(C\cdot C)$:

$$\mathcal{A}_4 = \int \frac{d^{2\times 4}C}{\mathrm{vol}(GL(2))}\, \frac{1}{(12)(23)} \delta^3(C\cdot C)\delta^{4|6}(C\cdot\Lambda). \tag{14.19}$$

With this, we can define on-shell diagrams for the ABJM theory just as for $\mathcal{N}=4$ by gluing together these basic 4-point vertices. Note that unlike for $\mathcal{N}=4$, n and k are not independent for ABJM: we always have $n=2k$.

It is easy to see that the on-shell representation of a BCFW shift is simply

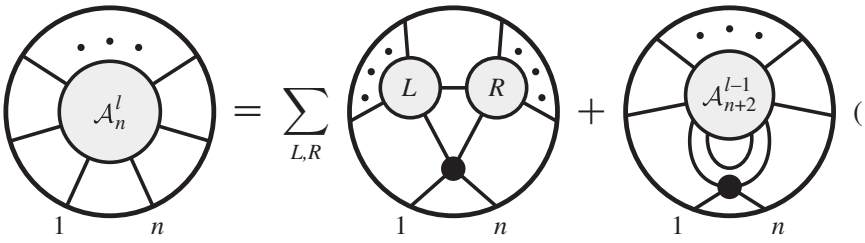


(14.20)

The action on the column vectors is simply a rotation between c_a and c_{a+1}:

$$c_a \mapsto c\,c_a - s\,c_{a+1}, \quad c_{a+1} \mapsto s\,c_a + c\,c_{a+1}\,. \tag{14.21}$$

And the all-loop integrand can be given in terms of on-shell diagrams just as before:



(14.22)

(for recent computations at one and two loops see [144–147]).

The rules for amalgamation are essentially identical to the $\mathcal{N}=4$ case—the only difference being some factors of i that must be included. In $(2+1)$ dimensions, because we write momenta as $p_a = \lambda_a \lambda_a$, switching $p_a \mapsto -p_a$ corresponds to taking $\lambda_a \mapsto i\lambda$. And so, when identifying two legs for the "projection" operation, instead of projecting relative to $(c_A - c_B)$, we must project relative to $(c_A - ic_B)$. The result is that minors of $C \in G(k,n)$ are related to those of the pre-image $\widehat{C} \in G(k+1, n+2)$ via

$$(a_1 \cdots a_k)|_C = (Aa_1 \cdots a_k)|_{\widehat{C}} + i\,(Ba_1 \cdots a_k)|_{\widehat{C}}\,. \tag{14.23}$$

It is very easy to see that, starting with elementary 4-point vertices in the null Grassmannian, amalgamation preserves this property; translated in terms of minors, this is the statement that for all a,

$$(c_1 \cdots c_{k-1} a)(d_1 \cdots d_{k-1} a) = 0. \tag{14.24}$$

This is trivial for the direct product. For projection, we easily verify that

$$\begin{aligned}
&(c_1 \cdots c_{k-1} a)(d_1 \cdots d_{k-1} a) \\
&\quad = [(Ac_1 \cdots c_{k-1} a) + i(Bc_1 \cdots c_{k-1} a)][(Ad_1 \cdots d_{k-1} a) + i(Bd_1 \cdots d_{k-1} a)] \\
&\quad = (Ac_1 \cdots c_{k-1} B)(Ad_1 \cdots d_{k-1} B) - (Bc_1 \cdots c_{k-1} A)(Ad_1 \cdots d_{k-1} B) \\
&\quad = 0.
\end{aligned}$$

Thus, amalgamation of many little null $G(2,4)$'s produces a point in the null Grassmannian $G(k,2k)$, together with the measure

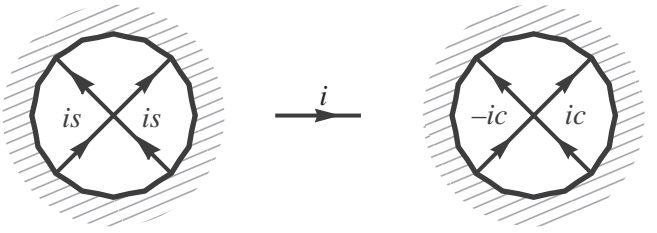
$$\prod_{\text{vertices } v} d\log(z_v). \tag{14.25}$$

Notice that an important difference between this and the case of $\mathcal{N}=4$ is that the fundamental variables are associated with the *vertices* of an on-shell graph, rather than its faces.

The measure on the top-cell can be given in terms of the C matrix via [141]

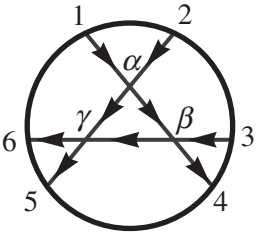
$$\frac{d^{k\times 2k}C}{\mathrm{vol}(GL(k))}\frac{\delta^{k(k+1)/2}\big(C\cdot C\big)}{(1\,2\cdots k)\cdots(k\,k+1\cdots 2k{-}1)}. \tag{14.26}$$

It is also straightforward to find the analog of boundary measurements by summing over all the paths joining sources to sinks in perfectly oriented graphs. We can orient each vertex with two incoming and two outgoing lines. Traversing any internal line contributes a factor of i, and at each vertex we get a is, ic, or $-ic$ according to:



(14.27)

As an example, consider the following 6-particle on-shell diagram:



(14.28)

We find that this diagram is associated with a configuration C in the Grassmannian represented by

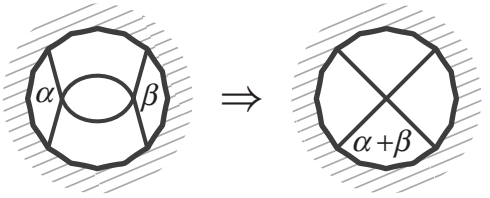
$$C=-i\left(\begin{array}{ccc|ccc} i & 0 & 0 & s_\alpha s_\beta & s_\alpha c_\beta c_\gamma + c_\alpha s_\gamma & c_\alpha c_\gamma - s_\alpha c_\beta s_\gamma \\ 0 & i & 0 & -c_\alpha s_\beta & s_\alpha s_\gamma - c_\alpha c_\beta c_\gamma & s_\alpha c_\gamma + c_\alpha c_\beta s_\gamma \\ 0 & 0 & i & c_\beta & -s_\beta c_\gamma & s_\beta s_\gamma \end{array}\right). \tag{14.29}$$

In general, we can write the $(k\times 2k)$ matrix representative $C\in G(k,2k)$ associated with any such graph in the form

$$C=-i\,(i1_{k\times k}|R_{k\times k}), \tag{14.30}$$

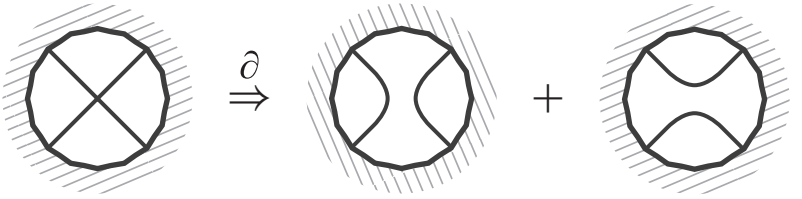
where R is an $SO(k)$-rotation matrix. This gives us a pretty interpretation for amalgamation. The basic 4-point vertex is just a rotation in two dimensions. Amalgamation provides a way of building general rotations in higher dimensions by a composing many rotations in two-dimensional subspaces. The 6-particle example above corresponds to a canonical way of representing three-dimensional rotations using Euler angles. The analog of the square move in ABJM looks much like the Yang–Baxter move, and represents the equality of two different Euler-angle representations of the same three-dimensional rotation.

Just as with $\mathcal{N}=4$, the invariant content of any reduced on-shell diagram is read off from its associated permutation. We also have an analog of reduction, looking at the 4-point bubble diagram connecting two 4-particle vertices with parameters α and β:



(14.31)

Finally, we can take a boundary, lowering the dimension by one, by deleting a vertex, and reconnecting the lines according to



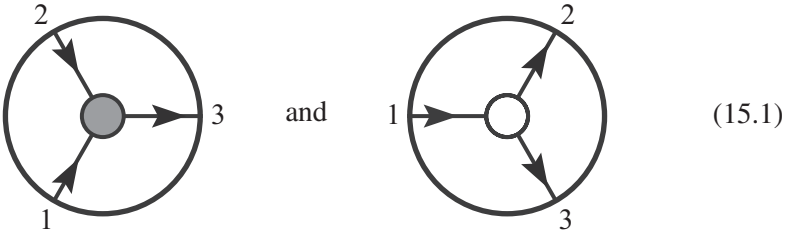
(14.32)

15

On-shell diagrams for theories with $\mathcal{N} < 4$ supersymmetries

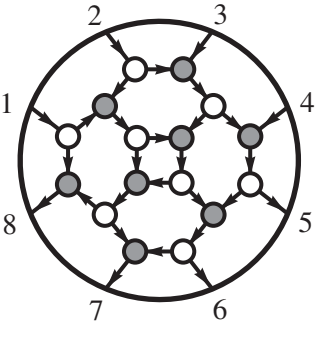
On-shell diagrams can be defined for any theory with fundamental trivalent vertices, and in particular for gauge theories with any number, $\mathcal{N}$, of supersymmetries. There is obviously a rich structure to be unearthed here; in this short section we will content ourselves with setting up some of the basic formalism and highlighting the central new mathematical object that makes an appearance—reflecting the physics of ultraviolet singularities—which are present in theories with less supersymmetry.

Let us begin our discussion by focusing on non-supersymmetric theories, those of "$\mathcal{N} = 0$". It is useful to represent the helicities involved in each basic three-particle vertex by giving each of the edges an orientation:



(15.1)

We can then glue these vertices together to build up more complex on-shell diagrams as before—leading to, for example:



(15.2)

In such decorated on-shell diagrams, the arrows are useful because they automatically encode the helicities of the internal particles involved. In general, we consider the particles as Grassmann-coherent states labeled by $\widetilde{\eta}^I$ for $I=1,\ldots,\mathcal{N}$. In theories with $\mathcal{N}<4$ supersymmetry, we have "+" and "−" multiplets, which include gluons of helicity ± 1 as their top components, respectively; thus, on-shell diagrams must be labeled in exactly the same way for any $\mathcal{N}<4$.

The Grassmannian formalism is just as powerful for integrating over the phase-space of the internal particles regardless of the amount of supersymmetry. However, when $\mathcal{N}<4$, the diagrams really are fundamentally oriented, whereas for $\mathcal{N}=4$ such an orientation merely encodes a convenient translation of the on-shell diagram into a particular gauge-fixed matrix-representative $C\in G(k,n)$. If the k incoming "source" indices are from a set A and the $(n-k)$ outgoing "sink" indices are from a, we find exactly the same linear relation between the external kinematical data:

$$\prod_A \delta^2\big(\widetilde{\lambda}_A - c_a^A\widetilde{\lambda}_a\big)\prod_A \delta^{\mathcal{N}}\big(\widetilde{\eta}_A - c_a^A\widetilde{\eta}_a\big)\prod_a \delta^2\big(\lambda_a + c_a^A\lambda_A\big), \tag{15.3}$$

where the c_a^A are exactly as in equation (4.58), which we reproduce below:

$$c_a^A = -\sum_{\Gamma\in\{A\rightsquigarrow a\}}\prod_{e\in\Gamma}\alpha_e. \tag{15.4}$$

The only difference between general $\mathcal{N}$ and $\mathcal{N}=4$ is the measure on the Grassmannian that ultimately encodes the on-shell differential form in terms of the auxiliary, Grassmannian degrees of freedom. For $\mathcal{N}=4$, we didn't have to include any Jacobian resulting from the elimination of internal variables, because the fermionic δ-functions always canceled the contributions between the internal bosons and internal fermions. However, when $\mathcal{N}<4$, these two factors do not cancel, and leave a net Jacobian contribution to the measure, which we may write as

$$\left(\prod_{\text{vertices } v}\frac{1}{\text{vol}(GL(1)_v)}\right)\left(\prod_{\text{edges } e}\frac{d\alpha_e}{\alpha_e}\right)\times\mathcal{J}^{\mathcal{N}-4}. \tag{15.5}$$

If the vertices of the graph are labeled i,j, then we define the *adjacency matrix* A_{ij} of the graph by

$$A_{ij} = \text{the weight of the } \textit{directed} \text{ edge } i\to j \text{ (if any)}; \tag{15.6}$$

then the Jacobian $\mathcal{J}$ is given by

$$\mathcal{J} = \det(1-A). \tag{15.7}$$

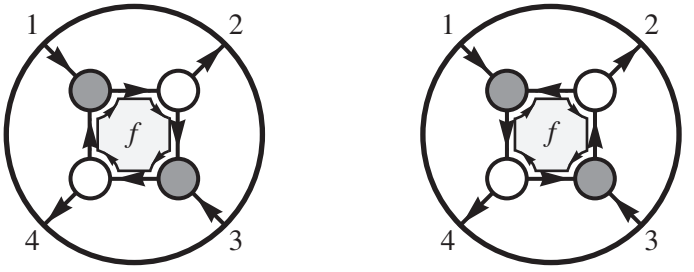
We know that the edge variables can only occur in the $GL(1)$ gauge-invariant "flux" combinations associated with faces, and we can give a simple formula for $\mathcal{J}$ in terms of these face variables. In general, if we have a collection of closed, *orientated* orbits bounding faces f_i, with *disjoint* pairs (f_i,f_j), *disjoint* triples (f_i,f_j,f_k), and so on, then $\mathcal{J}$ is given by:

$$\mathcal{J} = 1 + \sum_{\text{faces}} f_i + \sum_{\substack{\text{disjoint} \\ \text{pairs } i,j}} f_i f_j + \sum_{\substack{\text{disjoint} \\ \text{triples } i,j,k}} f_i f_j f_k + \cdots . \tag{15.8}$$

(Here, each 'face' is really a clockwise-oriented product of edge-variables around an orbit—and so may be the inverse of a face variable, or the product of face-variables that are bounded by a single orbit.)

For any oriented graph without any closed, oriented orbits, the spectrum is trivial, and $\mathcal{J} = 1$; for any such oriented on-shell diagram, the maximally supersymmetric and non-supersymmetric on-shell forms are identical. This is easy to understand because when an on-shell diagram is free of such oriented orbits, only gluons propagate internally (when the external states are gluons). In contrast, when there are oriented orbits, the rest of the super-multiplet *can* propagate internally, differentiating theories with different amounts of supersymmetry.

When an oriented on-shell diagram has closed, oriented orbits, the Jacobian is nontrivial. The simplest example occurs for four particles, where we can have



(15.9)

for which the corresponding Jacobians are

$$\mathcal{J}_A = 1 + f \qquad \text{and} \qquad \mathcal{J}_B = 1 + f^{-1}. \tag{15.10}$$

In order to compute the full on-shell process for fixed external sources and sinks, we have to sum over all the possible orientations of the internal graph. And so, in this case we would be obliged to sum over both diagrams, giving us a final contribution to the measure of

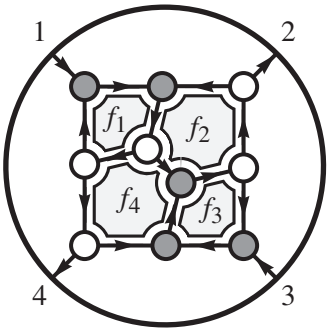
$$\mathcal{J}_A^{\mathcal{N}-4} + \mathcal{J}_B^{\mathcal{N}-4} = (1+f)^{\mathcal{N}-4} + (1+f^{-1})^{\mathcal{N}-4}. \tag{15.11}$$

Notice that when $\mathcal{N}=3$, the complete contribution is simply

$$\mathcal{J}_A^{-1}+\mathcal{J}_B^{-1}=(1+f)^{-1}+(1+f^{-1})^{-1}=1. \tag{15.12}$$

This is good, because the "+" and "−" super-multiplets of $\mathcal{N}=3$ combine to give us a complete $\mathcal{N}=4$ super-multiplet. Of course, when $\mathcal{N}<3$, the sum is not unity, and the result differs from what we would have found for $\mathcal{N}=4$.

Let us consider a somewhat more interesting example:



(15.13)

Here we have four closed orbits, and one disjoint pair of orbits: three of the orbits bound single faces, contributing f_1, f_2^{-1}, and f_3, and one orbit involving both f_2 and f_4 contributes $f_2^{-1}f_4^{-1}$; the disjoint pair are f_1, f_3. Putting everything together, we find that the complete Jacobian is:

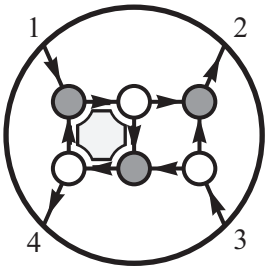
$$\mathcal{J}=1+(f_1+f_2^{-1}+f_3+f_2^{-1}f_4^{-1})+f_1f_3. \tag{15.14}$$

We stress again that the point in the Grassmannian obtained from amalgamation is the same as it is for the maximally supersymmetric theory; the *only* difference between the theories is the presence of the Jacobian factor $\mathcal{J}$ in the measure. The merge/unmerge moves still leave the point in the Grassmannian and the rest of the form invariant; but now, the square move and bubble-reduction—while leaving the point in the Grassmannian fixed—*can* change the measure.

If we consider a reduced graph with the dimension required to completely localize all the auxiliary variables associated with the matrix $C\in G(k,n)$, then the net effect is not particularly interesting—as theories with $\mathcal{N}<4$ differ from those with maximal supersymmetry only by the prefactor of $\mathcal{J}$ in the measure, evaluated at this particular point in $G(k,n)$. However, the situation is considerably more interesting when we consider on-shell graphs for which some auxiliary variables are not fixed by the δ-function constraints, leaving us with an integration measure over these internal degrees of freedom. Such graphs occur, for instance, in the forward-limits that generate loop integrands in the all-loop, on-shell BCFW

recursion (2.54). In such cases, the factor of $\mathcal{J}$ can lead to a qualitatively new set of singularities where poles are generated by $\mathcal{J}$.

As a simple example of such a situation, consider a "wrong" BCFW-bridge acting on the four-particle tree amplitude's on-shell graph:



(15.15)

The shift is "wrong" in the familiar sense of BCFW deforming the "wrong" helicities, for which the deformed amplitudes don't vanish at infinity. This is reflected in the on-shell graph by the presence of a closed oriented loop (making the resulting on-shell differential form differ for theories with different amounts of supersymmetry). Because this graph's measure includes the a nontrivial Jacobian $\mathcal{J}$, the corresponding function does not vanish in the deep ultraviolet—taking the shift-parameter $\alpha \to \infty$. This "pole at infinity" is characterized by the residue about $\mathcal{J} \mapsto 0$. Notice that this allows us to fully characterize the nontrivial, ultraviolet singularities present in theories with less than maximal supersymmetry. The presence of such poles indicate "lower-transcendentality" contributions to scattering amplitudes. For instance, the object above, (15.15), can be interpreted as the triple-cut of the one-loop four-particle amplitude, and the residue about $\mathcal{J} = 0$ computes the coefficient of the "triangle integral" for the amplitude. The coefficients of "bubbles" can be exposed in similar ways.

One of the most fundamental consequences of space-time locality is that the ultraviolet and infrared singularities are completely independent. It is fascinating to see that this physical fact is sharply captured by the Grassmannian formalism, where IR and UV singularities are associated with disparate contributions to the integration measure of the auxiliary Grassmannian: the positroid's "$d$log" measure captures all the long-distance singularities—where internal particles go on-shell—and the prefactor $\mathcal{J}$ captures ultraviolet singularities. This ultraviolet/infrared decoupling has an even more striking incarnation in the planar sector of the theory: it can be shown that $\mathcal{J}$ is completely regular everywhere in the positive part of $G(k,n)$—literally *separating* the ultraviolet singularities of $\mathcal{J}$ from infrared singularities of the positroid, their boundaries being completely disjoint!

16

Dual graphs and cluster algebras

So far in this book, we have extensively studied planar on-shell diagrams. In section 4.4, we introduced two natural classes of operations: *amalgamation*, the operation that allows us to build up very complex diagrams from very simple ones; and *mergers* and *square moves*, which allow us to connect very distinct on-shell diagrams that nevertheless encode the same physical information.

In this section we turn to the very obvious question that arises when dealing with planar diagrams of any sort: what are the corresponding dual graphs? what do they mean? and how are the operations we have found realized in terms of them? Of course, being two-colored, on-shell diagrams carry more information than ordinary graphs, and whatever definition of a dual graph we introduce must encode this additional information. Luckily, the theory of dual graphs for bipartite planar graphs is both known and simple; in fact, the dual of a *bipartite* graph is a familiar object in the physics of $\mathcal{N}=1$ supersymmetric gauge theories: it is a quiver diagram! Indeed, the connection between bipartite graphs and quiver gauge theories is already an active research area in the physics community and has led to beautiful constructions such as those described in [49–54]. Bipartite graphs are also intimately related to dimer models, with the recent mathematical work [41] particularly closely related to our discussion.

16.1 The 'dual' of an on-shell diagram

Recall that the dual of an ordinary planar graph (one without colored vertices) is obtained by drawing a vertex for each face, and connecting adjacent faces with edges. In our case, we have graphs on a *disc*, and so the faces of an on-shell diagram can be divided into two distinct classes: those in the *interior* of the graph, and those on the *exterior* (those adjacent to the boundary of the disc).

As mentioned above, the dual of a *bipartite* graph turns out to be none other than an *oriented* quiver diagram. Let us now describe how this dual "quiver" of

a general bipartite graph on a disc is defined. Let Γ denote a bipartite graph on a disc; we define a *flag* F of Γ to be the combination of one vertex of Γ with one edge connected to it. (Here, the word "flag" can just be taken as merely a name used in this construction; but—not surprisingly—this terminology stems from its more familiar use in algebraic geometry.) Each flag inherits a coloring according to the color of its vertex. We orient each black flag of Γ to be 'out of the vertex' and each white flag to be 'toward the vertex.' (Notice that every *external* edge is a member of a single flag, while each *internal* edge is part of two flags—one black, and one white.)

Let us suppose that Γ is *oriented* according to its internal flags—that is, we take each internal edge to be directed 'black-to-white.' Choosing a clockwise orientation for the boundary of each face f of Γ, we define the adjacency matrix for the dual graph $\widetilde{\Gamma}$ as follows: for each flag F, we define

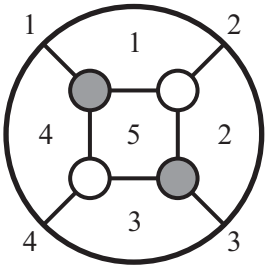
$$\delta^F_{f_1,f_2} \equiv \begin{cases} 0 & F \notin \left(\partial f_1 \bigcap \partial f_2\right); \\ +\frac{1}{2} & F \in \left(\partial f_1 \bigcap \partial f_2\right) \text{ and } F \text{ is oriented according to } \partial f_1; \\ -\frac{1}{2} & F \in \left(\partial f_1 \bigcap \partial f_2\right) \text{ and } F \text{ is oriented according to } \partial f_2. \end{cases}$$

The dual graph's adjacency matrix ε_{f_1,f_2} is then obtained by summing over all flags:

$$\varepsilon_{f_1,f_2} \equiv \sum_F \delta^F_{f_1,f_2}. \tag{16.1}$$

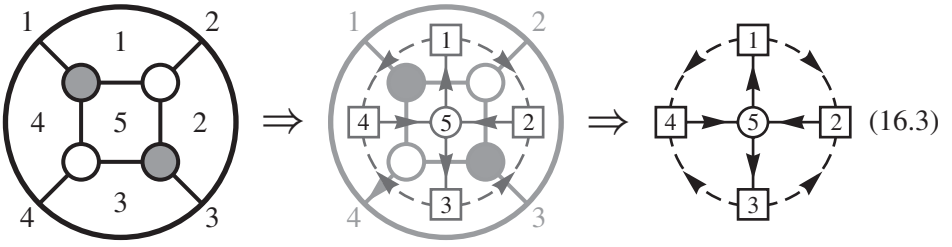
(The factors of "1/2" in the definition of $\delta^F_{f_1,f_2}$ are important for both mathematical and physical reasons, but will only show up in the final quiver for edges connecting external faces; to see this, recall that each internal edge is a member of *two* flags. They are important for the consistency of the general procedure of "amalgamation" [38], to be described more fully below.)

In order to illustrate the preceding discussion, let us consider—as usual—the on-shell diagram that generates the four-particle tree-amplitude,



(16.2)

Labeling the faces of the graph as indicated above, we find the dual quiver to be



(16.3)

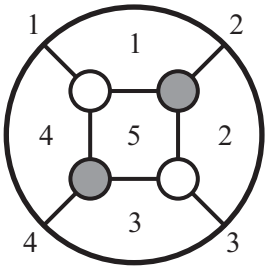
defined by the adjacency matrix $\varepsilon_{i,j}$:

$$\varepsilon = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & -1 \\ -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 1 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & -1 \\ -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 1 \\ 1 & -1 & 1 & -1 & 0 \end{pmatrix}. \tag{16.4}$$

Notice that in drawing (16.3), we have denoted each internal face by a circle and each external face by a square.

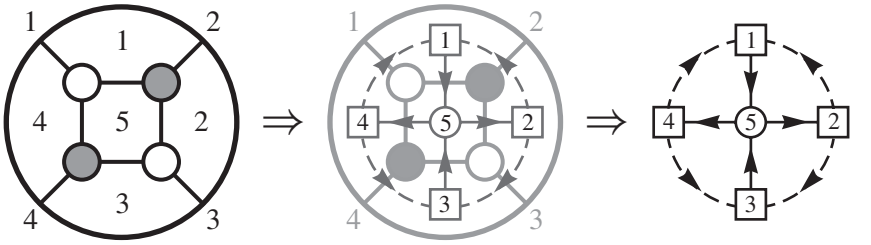
This quiver can be given a gauge theory interpretation. Let all the vertices represent $U(N)$ groups, with external ones being flavor groups and internal ones dynamical gauge groups. The adjacency matrix $\varepsilon_{i,j}$ denotes the number of bi-fundamental fields charged under the groups (i,j). Anomaly cancellation is the statement that all the rows of ε add up to zero. Of course, having a $1/2$ bi-fundamental field is clearly not possible in a physical theory; but a physical realization can be always be obtained by also including super-potential terms (see e.g. [53, 54]) which we will not need for our purposes of formal analogy.

Let's see how a square move affects the resulting dual-quiver; starting from



(16.5)

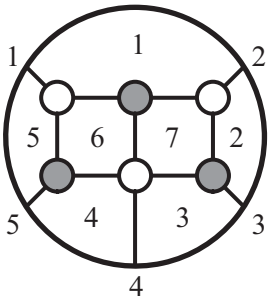
it is easy to find the corresponding quiver:



(16.6)

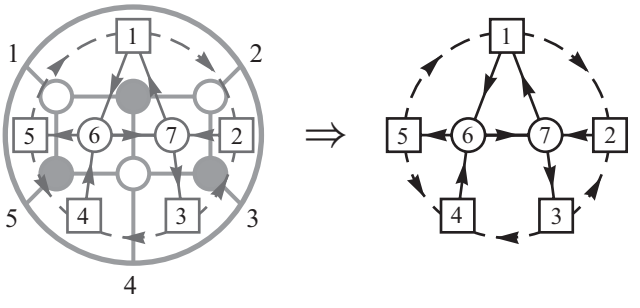
Very nicely, this new quiver corresponds to none other than the *Seiberg-dual* [4] of quiver (16.3) with respect to the internal node (5). In other words, the arrows connected to node (5) are reversed and new bi-fundamentals are created between two flavor nodes every time the bi-fundamentals connecting them to (5) can pair up. This happens when the arrows are in opposite directions. Whenever possible, the new bi-fundamentals pair up with existing ones to get a mass and disappear from the theory in the infrared. (Here, we again must stretch the analogy a bit, declaring that, e.g. the new bi-fundamental going from (1) to (2) pairs up with the existing 'half bi-fundamental' from (2) to (1) to leave behind a new 'half bi-fundamental' from (1) to (2).)

The factors of $1/2$ appearing in this discussion might seem like an unnecessary annoyance. In order to see the importance of these factors in the definition of $\delta^F_{f_1,f_2}$, let us consider the 5-particle on-shell diagram (in bipartite form)



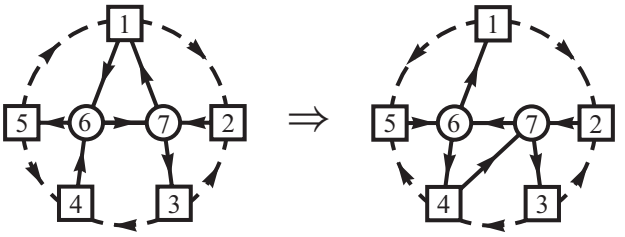
(16.7)

whose quiver is found to be given by



(16.8)

After performing a Seiberg duality on node (6), we find the quiver:

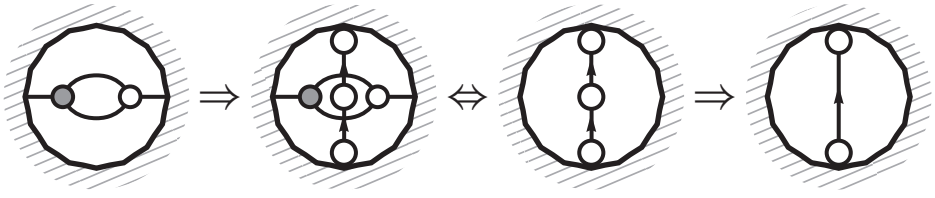


(16.9)

This can be seen to agree with the one obtained by expanding the 4-valent white vertex in the original on-shell diagram, applying a square move and merging the two new adjacent white vertices to make a new bipartite graph. This case shows that the factors of $1/2$ are needed in order to make the square move correspond to Seiberg duality when not all the four surrounding faces are external. As mentioned above, a fully physical alternative requires adding the natural super-potential terms associated with each closed loop in the quiver consistent with orientations (see e.g. [53, 54]).

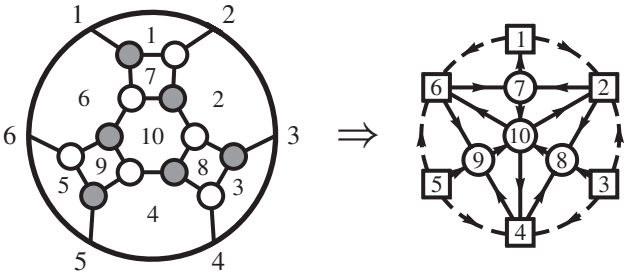
These examples make it clear that Seiberg duality on an internal nodes with *exactly* four edges corresponds to a square move in the on-shell diagram. Of course, once we have the dual-quiver of a given on-shell diagram is determined, it is tempting to perform Seiberg duality on any internal node—not only those of valency four. The question is then: what happens when a Seiberg duality is taken for nodes of arbitrary valency?

Let us start by considering the case of valency two. A valency-two node arises in the dual quiver only for on-shell diagrams with 'bubbles.' Recall that all nodes correspond to gauge groups $U(N)$; therefore, upon taking the Seiberg dual of a bivalent vertex, one finds that the rank goes from $N \mapsto N' = N - N = 0$—that is, it 'disappears' from the theory, leaving only the bi-fundamental created by the following duality:



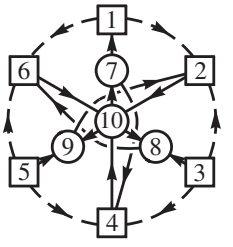
The resulting quiver is precisely that which would have been obtained for the on-shell diagram after bubble-deletion.

Let us now see the effect of applying Seiberg duality to a vertex whose valency is greater than four. Consider for example the dual-quiver of the following diagram for the top-cell of $G(3, 6)$:



(16.10)

Applying Seiberg duality to any of the nodes (7), (8), or (9), would correspond to a square move as expected. However, when we perform a Seiberg duality on the 6-valent vertex (10), we find the following:



(16.11)

This new quiver is nonplanar and therefore does not correspond to the dual of an on-shell diagram on the disc!

The presence of these new objects suggests that on-shell diagrams and operations like the square move are part of a much larger mathematical structure. As it turns out, the general story of such transformations is a very active area of research in mathematics: the theory of *cluster algebras*.

Let us now turn to a (basic) summary of this rich and more general story. The reason for doing so is that cluster algebras have recently made an appearance in the descriptions of seemingly unrelated developments in gauge theory and we anticipate that a deeper understanding will lead to important connections between these areas of physics and the physics of scattering amplitudes.

16.2 Cluster algebras: seeds, mutations, and cluster coordinates

The definition of a cluster algebra [35] begins with the notion of a *seed* and the various *mutations* that can take convert one seed into another. We will start with a general mathematical description of *seeds* and *mutations*, and try to relate these objects to on-shell diagrams as we proceed.

Seeds. A *seed* is a set of combinatorial data $\mathbf{s} = \{S, S_0, \varepsilon\}$, where S is a set, S_0 a distinguished subset of S called the *frozen subset*, and $\varepsilon_{i,j}$ is a skew-symmetric

matrix with $(i,j) \in S$ for which $\varepsilon_{i,j} \in \mathbb{Z}$ unless *both* of $(i,j) \in S_0$, in which case $\varepsilon_{i,j} \in \frac{1}{2}\mathbb{Z}$.

Notice that for the discussion about the dual-quiver of an on-shell diagram is an example of a seed where S is the set of faces, S_0 the subset of faces on the boundary of the disc, and ε precisely as defined above. Notice that the distinction between S and S_0 nicely matches the physical distinction between flavor and dynamical groups of the quiver theory.

Mutations. Given a seed $\mathbf{s} = \{S, S_0, \varepsilon\}$ and any non-frozen element $k \in (S \setminus S_0)$, the *mutation* of $\mathbf{s}$ "in the direction k" is $\mu_k(\mathbf{s}) \equiv \{S, S_0, \varepsilon'\}$ where the *mutated* matrix ε' is given by the Fomin Zelevinsky formula [35]:

$$\varepsilon'_{i,j} \equiv \begin{cases} -\varepsilon_{i,j} & k \in \{i,j\}; \\ \varepsilon_{i,j} & k \notin \{i,j\} \text{ and } \varepsilon_{i,k}\varepsilon_{k,j} \leq 0; \\ \varepsilon_{i,j} + |\varepsilon_{i,k}| \cdot \varepsilon_{k,j} & k \notin \{i,j\} \text{ and } \varepsilon_{i,k}\varepsilon_{k,j} > 0. \end{cases} \tag{16.12}$$

Notice that this procedure is *involutive*: the mutation of ε' in the direction k is the original matrix ε. Moreover, the rule for mutations *exactly* corresponds to the transformation of bi-fundamental matter fields occurring under a Seiberg duality of a dynamical gauge-group node (corresponding to an element of $(S \setminus S_0)$).

Cluster coordinates. Recall the description the *face variables* $\{f_i\}$ associated with any on-shell diagrams given in section 4.5; one of the most important features of the face variables is that if they are chosen to be *positive* (see section 5.4) then they remain positive after a square move. Also important is that face variables are *canonical* coordinates on the Grassmannian configuration C associated with an on-shell diagram: that is, the positroid volume-form on the Grassmannian—when expressed in these coordinates—is simply

$$d\log(\widehat{f}_1) \wedge d\log(\widehat{f}_2) \wedge \cdots \wedge d\log(\widehat{f}_d), \tag{16.13}$$

where we have defined the *rescaled* coordinates $\widehat{f}_i \equiv f_i/f_0$. Importantly, this measure is invariant (up to a sign) under square moves.

Now that we allow mutations that are *not* square moves and obtain quivers that are not related to planar on-shell diagrams, one could ask if there is any analog of 'face variables' or the volume form on the Grassmannian. It turns out that in fact there are two sets of coordinates that can be defined in general, and one of them coincides with the face variables described in section 4.5 when restricted to on-shell diagrams. These are known as the cluster $\mathcal{A}$-coordinates and the cluster $\mathcal{X}$-coordinates.

Given a seed $\mathbf{s}$ the two sets of coordinates are each parameterized by the set S—the set of 'faces' in the case of planar on-shell diagrams. Let us denote the $\mathcal{X}$-coordinates by $\{X_i\}$ and the $\mathcal{A}$-coordinates by $\{A_i\}$. (This is a slight abuse of notation because we have suppressed their dependence on the seed $\mathbf{s}$.) *A priori*, one makes a proposal for what the coordinates are in terms of the data defining the problem at hand, and then the kind of variables will be determined by their behavior under mutations.

Given a mutation of seeds $\mu_k : \mathbf{s} \mapsto \mathbf{s}'$, the cluster coordinates assigned to these seeds are related as follows. If we denote the cluster coordinates related to the seed $\mathbf{s}$ by X_i and A_i, and the ones assigned to the seed $\mathbf{s}'$ by X'_i and A'_i, then we have mutation formulae for $\mathcal{A}$-coordinates:

$$A_k A'_k \equiv \prod_{j|\varepsilon_{kj}>0} A_j^{\varepsilon_{kj}} + \prod_{j|\varepsilon_{kj}<0} A_j^{-\varepsilon_{kj}}, \quad \text{and} \quad A'_i = A_i \text{ for } i \neq k; \tag{16.14}$$

and for $\mathcal{X}$-coordinates we have

$$X'_i \equiv \begin{cases} X_k^{-1} & i = k; \\ X_i \left(1 + X_k^{-\mathrm{sgn}(\varepsilon_{ik})}\right)^{-\varepsilon_{ik}} & i \neq k. \end{cases} \tag{16.15}$$

(In (16.14), if any of the sets $\{j|\varepsilon_{kj} > 0\}$ or $\{j|\varepsilon_{kj} < 0\}$ is empty, the corresponding monomial is 1.)

The set of transformations among cluster coordinates obtained by composing mutations are known as *cluster transformations*.

Cluster $\mathcal{A}$-coordinates and mutation formulae are the main ingredients of the definition of cluster algebras of Fomin and Zelevinsky [35]. So far in this book these coordinates have not made an appearance but they can be nicely defined for on-shell diagrams using left-right paths as follows. For each left-right path take the end point label and write it on all faces to the left of the path. If the on-shell diagram is reduced and represents a cell of $G(k,n)$, this procedure provides each face with k labels. The Plücker coordinates given by the sets of labels in each face provides a set of cluster $\mathcal{A}$-coordinates. These coordinates are generically non-vanishing and, under a square move, indeed transform as $\mathcal{A}$-coordinates.

Cluster $\mathcal{X}$-coordinates and mutation formulae (16.15) describe a dual object, introduced in [37] under the name *cluster $\mathcal{X}$-variety* or *cluster Poisson variety*. The transformation (16.15) is precisely the way face variables were found to transform in section 4.6. So, our face variables are an example of cluster $\mathcal{X}$-coordinates.

One of the key features of cluster $\mathcal{X}$-varieties is that they are endowed with a natural Poisson structure. We can define

$$\{X_i, X_j\} \equiv \varepsilon_{ij} X_i X_j. \tag{16.16}$$

Of course, the crucial fact is that the cluster transformations (16.15) preserve the Poisson bracket. This fascinating structure has made an appearance in a number of physical settings, but has not yet played a role in our understanding of scattering amplitudes.

Since the cluster transformations are given by subtraction free formulas, they identify the sets of points with the real positive cluster coordinates assigned to different seeds. Gluing these sets according to the cluster transformations, we arrive at the spaces $\mathcal{A}^+$ and respectively $\mathcal{X}^+$ on which all cluster coordinates are perfectly well defined and take positive real values.

The Grassmannian is an example of a cluster $\mathcal{X}$-variety, and the corresponding space of positive points $\mathcal{X}^+$ is nothing else but the positive Grassmannian.

It turns out that just as for face variables in on-shell diagrams, one can define canonical volume forms for both kinds of coordinates [46]. Given a seed $\mathbf{s}$, consider the volume forms

$$\text{vol}^{\mathbf{s}}_{\mathcal{A}} \equiv d\log A_1 \wedge \cdots \wedge d\log A_n, \qquad \text{vol}^{\mathbf{s}}_{\mathcal{X}} \equiv d\log X_1 \wedge \cdots \wedge d\log X_n. \tag{16.17}$$

It is an easy but fundamental fact that the cluster transformations preserve the corresponding volume form up to a sign. Precisely, given a seed mutation $\mathbf{s} \mapsto \mathbf{s}'$, we have

$$\text{vol}^{\mathbf{s}'}_{\mathcal{A}} = -\text{vol}^{\mathbf{s}}_{\mathcal{A}}, \qquad \text{vol}^{\mathbf{s}'}_{\mathcal{X}} = -\text{vol}^{\mathbf{s}}_{\mathcal{X}}. \tag{16.18}$$

To check the first identity, let us do a mutation at k. Then only the coordinate A_k changes, and due to the exchange relation (16.14), one has

$$d\log A'_k + d\log A_k = 0 \mod dA_j, \quad \text{where } j \neq k. \tag{16.19}$$

To check the second, notice that under a mutation at k, one has $d\log X'_k = -d\log X_k$, while $d\log X'_j = d\log X_j$ modulo dX_k. These forms are known as the $\mathcal{A}$- and $\mathcal{X}$-cluster volume forms. Of course, our top-form on the positive Grassmannian precisely coincides with the $\mathcal{X}$ volume form.

Singularities of the cluster volume-form and frozen variables

Take a variety equipped with a cluster $\mathcal{A}$-coordinate system $\{A_i\}$. Let us assume that $k \in (S \setminus S_0)$ is non-frozen, and $\varepsilon_{kj} \neq 0$ for some j. Then the residue of the cluster volume form $\text{vol}_{\mathcal{A}}$ at the locus $A_k = 0$ is zero:

$$\text{Res}_{A_k=0}(\text{vol}_{\mathcal{A}}) = 0. \tag{16.20}$$

Indeed, the residue is given by $\text{Res}_{A_k=0}(\text{vol}_{\mathcal{A}}) = \pm \bigwedge_{i \neq k} d\log A_i$. Since k is non-frozen, there is an exchange relation (16.14). It implies a monomial relation

on the locus $A_k = 0$:

$$\prod_j A_j^{\varepsilon_{kj}} = -1. \tag{16.21}$$

Since ε_{kj} is not identically zero, the monomial in (16.21) is nontrivial. This implies that the form $\bigwedge_{i\neq k} d\log A_i$ vanishes at the $A_k = 0$ locus.

This explains the role of frozen variables in a cluster coordinate system $\{A_i\}$ on a space M. Indeed, a coordinate A_k, with $\varepsilon_{kj} \neq 0$ for some j, can be declared non-frozen only if $\mathrm{Res}_{A_k=0}(d\log A_1 \wedge \cdots \wedge d\log A_n) = 0$. This condition just means that the functions $A_1, \ldots, \widehat{A_k}, \ldots, A_n$ become dependent on every component of the $A_k = 0$ locus.

16.3 Cluster amalgamation

The "atomic" principle, in which complicated objects and their properties can all be simply derived from constituent building blocks, has played a fundamental role in understanding on-shell diagrams, and is more generally making an appearance more and more often in both physics and mathematics. Given the potential importance of this phenomenon, let us now describe the more general procedure of amalgamation of cluster structures, of which our sense of amalgamation is a special case. We find it convenient to use a different but equivalent description of seeds which is known as the geometric description. A definition of amalgamation via the combinatorial description of seeds is given in [38].

The following geometric definition is taken from [37]:

Definition: A *seed* is a set of combinatorial data $\mathbf{s} = \{\Lambda, \Lambda_0, \{e_i\}, \varepsilon\}$, where Λ is a free abelian group, Λ_0 is a distinguished subgroup of Λ, $\{e_i\}$ is a basis of Λ such that Λ_0 is generated by a subset of *frozen* basis vectors, and $\varepsilon_{i,j} \equiv \varepsilon(e_i, e_j)$ is a skew-symmetric bilinear form on Λ such that $\varepsilon_{i,j} \in \mathbb{Z}$ unless *both* of $(e_i, e_j) \in \Lambda_0$, in which case $\varepsilon_{i,j} \in \frac{1}{2}\mathbb{Z}$.

To see that this definition of a seed is equivalent to the previous one, note that given combinatorial data $\{S, S_0, \varepsilon\}$, the abelian group Λ is the free abelian group generated by the set S, where the generator e_i is the one assigned to an element $i \in S$. The subgroup Λ_0 is then generated by the subset S_0, and the bilinear form $\varepsilon(\cdot,\cdot)$ is defined as above. Vice versa, given data $\{\Lambda, \Lambda_0, \{e_i\}, \varepsilon\}$, the set S is the set parameterizing the basis vectors, and so on.

Given this geometric definition, it is a good point to mention that mutations can be interpreted as half-reflections. This brings us again closer to the known description of Seiberg duality in quiver gauge theories as Weyl reflections where coupling constants in the form $1/g_i^2$ transform as root vectors e_i. In full generality we have that the seed $\mathbf{s}'$ obtained from $\mathbf{s}$ by the mutation in the direction k is defined

by changing the basis $\{e_i\}$ (the rest of the data stays the same). The new basis $\{e'_i\}$ is defined as a half-reflection of the one $\{e_i\}$ along the hyperplane $\varepsilon(e_k,\cdot)=0$:

$$e'_i \equiv \begin{cases} e_i + [\varepsilon_{ik}]_+ e_k & \text{if} \quad i \neq k \\ -e_k & \text{if} \quad i = k. \end{cases} \tag{16.22}$$

Here we set $[\alpha]_+ \equiv \alpha$ if $\alpha \geq 0$ and $[\alpha]_+ \equiv 0$ otherwise. One can check that formula (16.22) amounts to formula (16.12), telling how the ε-matrix changes under mutations.

By definition, the frozen/non-frozen basis vectors of the mutated seed are the images of the frozen/non-frozen basis vectors of the original seed. The composition of two mutations in the same direction k is no longer the identity, but rather an isomorphism of seeds.

We are now ready to turn to the amalgamation procedure. Take a pair of seeds, where we emphasize the set of frozen basis vectors $\{f_i\}$:

$$\mathbf{s}' = \Big\{\Lambda', \varepsilon', \{e'_i\}, \{f'_j\}\Big\}, \qquad \mathbf{s}'' = \Big\{\Lambda'', \varepsilon'', \{e''_i\}, \{f''_j\}\Big\}. \tag{16.23}$$

First, we define their direct product as follows:

$$\mathbf{s}' \otimes \mathbf{s}'' \equiv \Big\{\Lambda, \varepsilon, \{e_i\}, \{f_j\}\Big\}, \tag{16.24}$$

where $\Lambda \equiv \Lambda' \oplus \Lambda''$, and the form ε is defined to be the direct product of the forms ε' and ε''. The basis vectors and the frozen ones, are inherited in an obvious way.

Next, given a seed $\mathbf{s} = \Big\{\Lambda, \varepsilon, \{e_i\}, \{f_j\}\Big\}$, and a pair of frozen basis vectors f_a and f_b, we define the *reduced seed*

$$\mathbf{s}_{a*b} = \Big\{\Lambda_{a*b}, \varepsilon_{a*b}, \{e_s\}, \{f_t\}\Big\}. \tag{16.25}$$

Here Λ_{a*b} is a co-rank-one subgroup of Λ whose basis vectors are $(f_a + f_b)$ and those of Λ different from f_a and f_b, and the vectors $(f_a + f_b)$ and $f_j \notin \{f_a, f_b\}$ are the frozen ones; the form ε_{a*b} is the restriction to Λ_{a*b} of the form ε on Λ.

Given a pair of seeds (16.23), a pair of subsets $\{f'_a\}, a \in A$ and $\{f''_b\}, b \in B$ of the frozen basis vectors in $\mathbf{s}'$ and $\mathbf{s}''$, and an isomorphism of sets $\varphi : A \to B$, we define the *amalgamation* $\mathbf{s}' *_\varphi \mathbf{s}''$ of the seeds $\mathbf{s}'$ and $\mathbf{s}''$ along φ. This is done in a few steps:

1. take the direct product $\mathbf{s}' \otimes \mathbf{s}''$;
2. perform the reduction along a pair of frozen vectors $f'_a, f''_{\varphi(a)}$, for each $a \in A$;
3. if the restriction of the form ε of the seed $\mathbf{s}' *_\varphi \mathbf{s}''$ to the basis vectors $f'_a + f''_{\varphi(a)}$ with $a \in A$ takes values in $\mathbb{Z}$, then *defrost* these basis vectors by declaring them to be *unfrozen* (meaning that we now allow mutations at these vectors).

The first two steps amounts to taking the subgroup of $\Lambda' \oplus \Lambda''$ generated by the vectors $f'_a + f''_{\varphi(a)}$, $a \in A$, and the rest of the basis vectors, and inducing on it a seed structure. The amalgamation of seeds evidently commutes with the seed mutations.

The cluster coordinates X_i on the set S are related to the ones X'_i and X''_i by

$$X_i \equiv \begin{cases} X'_i & i \in (S' \setminus A); \\ X''_i & i \in (S'' \setminus \varphi(A)); \\ X'_a X''_{\varphi(a)} & i = a \in A. \end{cases} \tag{16.26}$$

It is easy to check that the amalgamation respects the Poisson structure. For the cluster $\mathcal{A}$-coordinates, we have

$$A_i \equiv \begin{cases} A'_i & i \in (S' \setminus A); \\ A''_i & i \in (S'' \setminus \varphi(A)); \\ A'_a \big(= A''_{\varphi(a)} \big) & i = a \in A. \end{cases} \tag{16.27}$$

The algebra generated by the cluster $\mathcal{X}$-coordinates of the amalgamated seed embeds to the product of similar algebras assigned to the original seeds via formulae (16.26).

Contrary to this, the algebra generated by the cluster $\mathcal{A}$-coordinates of the amalgamated seed is the quotient of the product of the similar algebras assigned to the original seeds: we impose the relations $A'_a = A''_{\varphi(a)}$.

16.4 (Brief) overview of cluster structures in physics

The theory of cluster algebras had its origins in a very unexpected area: the study of totally positive square matrices. This investigation began in the 1930s with Gantmacher and Krein [43] and Schoenberg [44], and had immediate applications to the theory of oscillators in classical mechanics.

The notion of total positivity was vastly generalized by Lusztig [34] to the case of arbitrary split real reductive groups G. Lusztig defined the positive part of group $G_{>0}$ by using the Chevalley generators. The study of total positivity, related parameterizations and canonical bases in simple Lie groups theory led to discovery of cluster algebras [35].

A crucial feature of cluster Poisson varieties in connection with physics is that they provide a very general example of *non-perturbative* quantization [46].

Recall that to quantize a Poisson space $(X, \{,\})$ means to deform its algebra of functions to a noncommutative algebra $\mathcal{O}_g(X)$, depending on a "coupling" constant $g > 0$ (normally referred to in the literature as "$\hbar$") so that $\widehat{a}\widehat{b} - \widehat{b}\widehat{a} = g\{a,b\} + \ldots$, and represent the algebra $\mathcal{O}_g(X)$ by operators in a Hilbert space. The Heisenberg quantization does this for a flat space with canonical coordinates (p_i, q_i). Kontsevich [148] proved that a perturbative version of the algebra $\mathcal{O}_g(X)$, where the dependence on the coupling g is via formal power series, always exists.

Any cluster Poisson variety $\mathcal{X}$ admits a non-perturbative quantization, which is manifestly invariant under the "S-duality" $g \to g^{-1}$. It comes with a series of $*$-representations in Hilbert spaces, modeled on a single Hilbert space $L^2(\mathcal{A}^+, \Omega_{\mathcal{A}})$, and defined using the space $\mathcal{A}^+$ of positive real points of the dual cluster variety $\mathcal{A}$, with the canonical cluster volume form $\Omega_{\mathcal{A}}$ providing the Lebesgue measure there.

Many (if not most) interesting examples of cluster structures appear when one couples a reductive Lie group G to a topological surface S, studying moduli spaces of flat G-bundles on a topological surface S and related moduli spaces introduced and studies in [36]. The corresponding spaces of positive real points are the Higher Teichmüller spaces related to the pair (G, S).

Let us give a broad perspective description of how the non-perturbative cluster quantization, and in particular the canonical cluster volume form, play a crucial role in defining a Hilbert space that has now made an appearance several times in quantum field theory.

The staring point is a *decorated surface* S. Let S be a surface with or without boundary, and a finite collection of points on the boundary, considered modulo isotopy. For example, a disc with n points on the boundary is the topological background for the n-particle scattering amplitudes.

Given S and a split reductive Lie group G, there are two moduli spaces defined in [36] closely related to the moduli space of G-bundles with flat connections on S:

$$\mathcal{X}_{G,S} \quad \text{and} \quad \mathcal{A}_{G,S}. \tag{16.28}$$

The first is equipped with an $\mathcal{X}$-cluster atlas, and the second with an $\mathcal{A}$-cluster atlas.

This immediately implies that the sets $\mathcal{X}^+_{G,S}$ and $\mathcal{A}^+_{G,S}$ of real positive points of these spaces are defined. This is the dual pair of higher Teichmüller spaces assigned to (G, S).

As described in [37], the existence of the cluster atlas on the space $\mathcal{X}_{G,S}$ implies that the algebra $\mathcal{O}(\mathcal{X}_{G,S})$ of regular functions on this space admits a canonical noncommutative q-deformation to a $*$-algebra $\mathcal{O}_q(\mathcal{X}_{G,S})$, where $q = \exp(\pi i g)$ (for $g > 0$). It is invariant under the action of the mapping class group of S.

The *modular double* of the algebra $\mathcal{O}_q(\mathcal{X}_{G,S})$ is the tensor product of the original $*$-algebra with the coupling g, and the $*$-algebra related to the Langlands dual group G^L at the "inverse" $1/(d_G g)$ of the coupling ($d_G = 1$ for simply-laced group G):

$$\mathcal{O}_q(\mathcal{X}_{G,S}) \otimes \mathcal{O}_{q^\vee}(\mathcal{X}_{G^L,S}), \tag{16.29}$$

where $q = \exp(\pi i g)$ and $q^\vee = \exp(\pi i / d_G g)$.

It follows from the general result on quantization of cluster varieties proved in [46] that the modular double admits a series of $*$-representations in Hilbert

spaces $\mathcal{H}_\chi$, depending on a parameter $\chi \in \mathbb{R}^m$. Here m is the dimension of the center of the Poisson bracket on the space $\mathcal{X}_{G,S}$, and χ is a unitary character of the center of the algebra $\mathcal{O}_q(\mathcal{X}_{G,S})$. The Hilbert spaces $\mathcal{H}_\chi$ are expected to be the spaces of conformal blocks for higher Toda theories.

On the other hand, since the space $\mathcal{A}_{G,S}$ has a cluster $\mathcal{A}$-atlas, it carries the canonical volume form $\Omega_{\mathcal{A}}$. The latter restricts to a canonical volume form on the positive real space $\mathcal{A}^+_{G,S}$. Therefore we arrive at a canonical Hilbert space assigned to the pair (G,S):

$$L^2(\mathcal{A}^+_{G,S}, \Omega_{\mathcal{A}}). \tag{16.30}$$

The mapping class group of S acts by its unitary symmetries. It was proved in [46] that this Hilbert space is the integral of the spaces of operators acting in the Hilbert spaces $\mathcal{H}_\chi$:

$$L^2(\mathcal{A}^+_{G,S}, \Omega_{\mathcal{A}}) = \int \mathcal{H}^*_\chi \otimes \mathcal{H}_\chi d\chi. \tag{16.31}$$

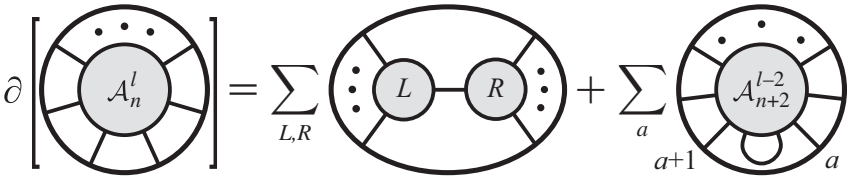
We conclude that the positive structure and the canonical cluster volume form on the space $\mathcal{A}_{G,S}$ provide us with the Hilbert space $L^2(\mathcal{A}^+_{G,S}, \Omega_{\mathcal{A}})$ describing the conformal blocks.

As we have seen in this book, it is quite amazing that exactly the same data—the positive structure and the canonical cluster volume form—for the Grassmannian $G(k,n)$ also provide us the measure for the integrand of the scattering amplitudes in $\mathcal{N}=4$. It is even more striking that there are structures crucial for each of these stories that have not made an appearance in the other: we need the quantized dual $\mathcal{X}$-moduli space in one, and the rich external kinematic data in the other. This strongly suggests that a deeper study is bound to reveal the roles of the "missing structures" in each of the stories and lead to a beautiful unified picture.

17

On-shell representations of scattering amplitudes

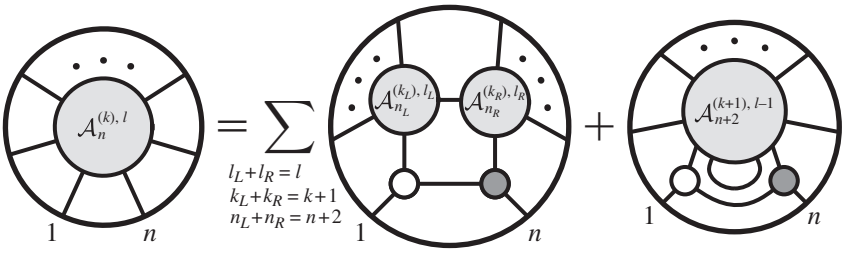
Although we have focused on understanding individual on-shell diagrams for most of the book, let us return to a study of how these can combine to entire scattering amplitudes. As discussed in Chapter 2, the defining property of the full amplitude is that it satisfies the "differential equation"



(17.1)

which specifies the two kinds of singularities it can have—corresponding to "factorization channels" and "forward-limits," respectively. All known representations of scattering amplitudes can be thought of as particular ways of building objects with these—and *only* these—(co-dimension one) singularities.

The usual Feynman-diagrammatic expansion for scattering amplitudes makes these singularities (together with conformal invariance) manifest, but at the cost of introducing unphysical, off-shell variables and (gauge-)redundancies that obscure the underlying Yangian invariance of the theory. (The same can be said for the equivalent Wilson-loop representation—except that it is the *dual* conformal symmetry that is made manifest.) By contrast, the BCFW recursion relations,



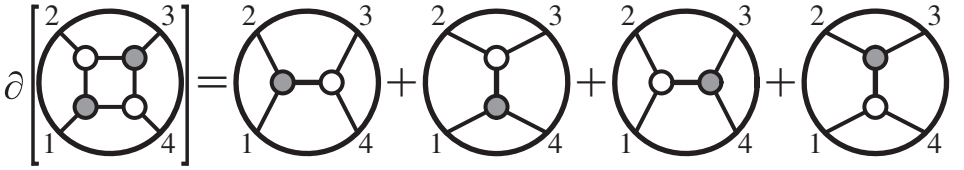
(17.2)

can be understood as a *direct* integration of the defining equation (17.1), and provides us with a representation of scattering amplitudes for which *every term* enjoys the full Yangian invariance of the theory. However, the recursion requires that two legs be singled out to play a special role—in (17.2), these are the legs $(n\,1)$. Although this choice is arbitrary, it breaks the cyclic symmetry of the complete amplitude, and makes manifest only a rather small subset of the singularities required by (17.1).

Of course, the BCFW recursion relations can be derived from field theory, starting with either the "scattering amplitude" [14] or "Wilson loop" [149–151] picture (for the relation to light-like correlation functions, see [152–156]). We will however begin by showing how they can also be proven directly by induction. That is, we will show that the boundary of (17.2) includes *precisely* the singularities required by (17.1); this proof will be entirely *diagrammatic*. In section 17.2 we will review some important features encountered in the tree-level ($l=0$) version of the recursion relations, and in section 17.3 we will see how the structure of tree amplitudes is reflected at loop level, giving rise to a canonical—purely '$d$log'—form for all loop-integrands.

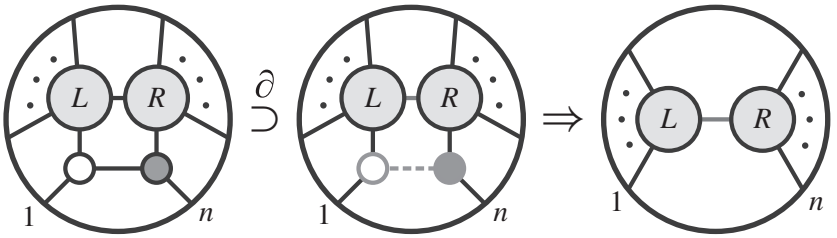
17.1 (Diagrammatic) proof of the BCFW recursion relations

Let us take the BCFW recursion relations, (17.2) as an ansatz, and demonstrate inductively that its boundary includes *all* the correct factorization channels and forward-limits, and no other singularities (for earlier work along these lines see [157, 158]). Recall that the four-point tree-amplitude, $\mathcal{A}_4^{(2),l=0}$, manifestly has all the correct factorization channels in its boundary:

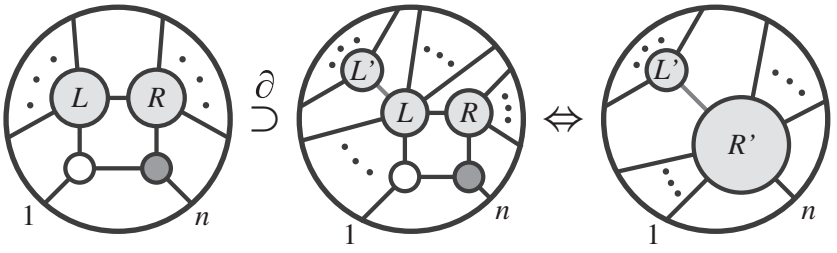


We may therefore suppose that the ansatz is correct for all amplitudes $\mathcal{A}_{\widehat{n}}^{(\widehat{k}),\widehat{l}}$ with $\widehat{n}<n$, $\widehat{k}\leq k$, and $\widehat{l}\leq l$; we must show that this suffices to prove that it also holds for $\mathcal{A}_n^{(k),\ell}$. We may divide the argument into two parts: first, demonstrating that the boundary includes all the correct factorization channels; and then showing that it includes all the correct forward-limits.

Among the factorization channels, those for which particles 1 and n are on opposite sides are trivially present:

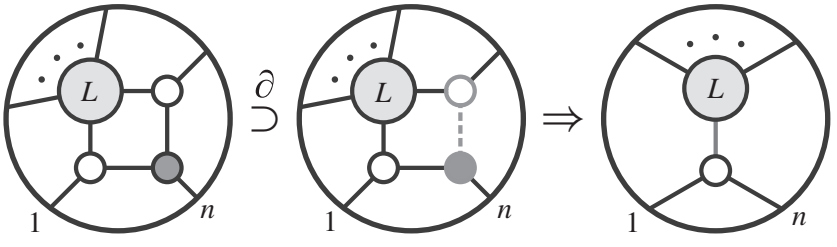


What we first need to check is that the BCFW recursion formula also generates all those factorizations for which 1 and n are on the same side. Factorization channels for which legs 1 and n are not alone on one side arise from the factorizations of the bridged amplitudes. For example, the boundaries of the left-amplitudes include

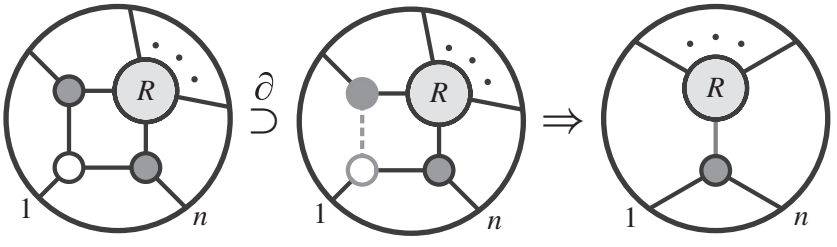


where we have used our induction hypothesis to identify the terms appearing on the right side of the factorization as a lower-point amplitude denoted R'. We also have the analogous diagrams arising from the right-amplitudes.

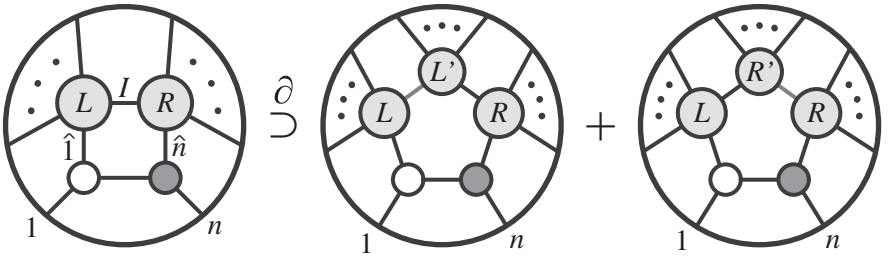
The case of a two-particle factorization involving *just* 1 and n together, however, arises somewhat differently. The factorization for which particles 1 and n are connected via a $\mathcal{A}_3^{(1)}$-vertex arises from the boundary



Similarly, the case where particles 1 and n are connected via a $\mathcal{A}_3^{(2)}$-vertex arises from

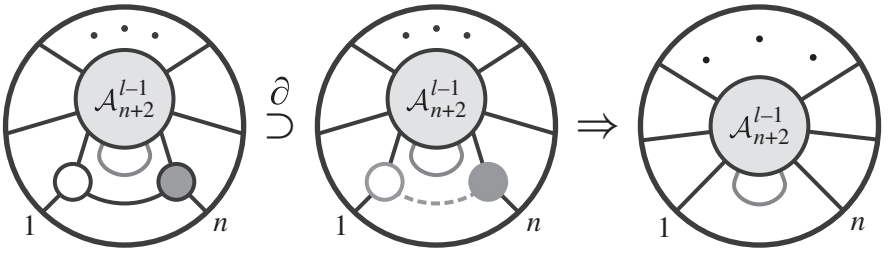


We have therefore shown that all factorization channels are present in the boundary of the BCFW ansatz. However, we must also show that these are the *only* such boundaries. Our induction hypothesis would suggest that such 'spurious' poles could arise from factorizations of separating $(\widehat{1}I)$ on the left, or $(I\widehat{n})$ on the right:

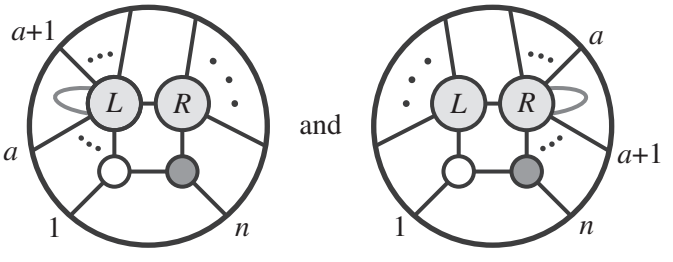


Conveniently, such boundaries are *always* generated symmetrically from the left- and right-amplitudes, and cancel in the sum.

Let us now demonstrate that the BCFW recursion ansatz generates all the correct forward-limits as codimension-one boundaries—and *only* these. As with the factorization channels, the BCFW recursion ansatz always makes one of the forward-limits manifest—those where the forward-limit is taken between 1 and n:

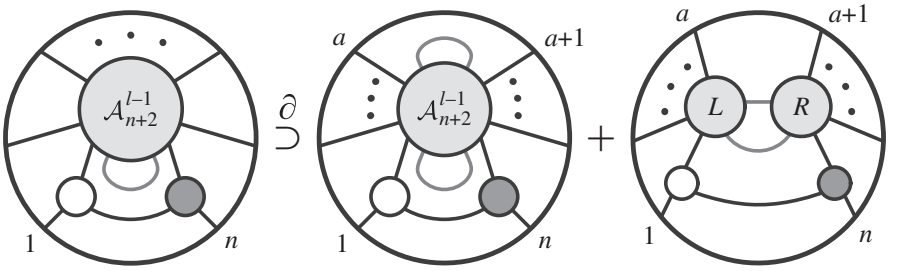


When the identified legs are not between $(n\,1)$, but say $(a\,a+1)$, something more interesting happens. Some of these arise trivially from the boundary of 'bridged' terms in the recursion:

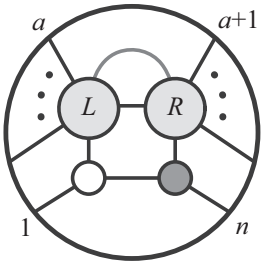


but these terms alone do not represent the complete BCFW-representation of the lower-loop, higher-point amplitude including the identified legs: the problem is that we are missing both the terms where the identified legs (before the forward-limit) are separated across the BCFW-bridge, and also the terms for which they are identified in the 'forward-limit' term. By our induction hypothesis, both of these terms arise from the boundary of the forward-limit term, as factorization

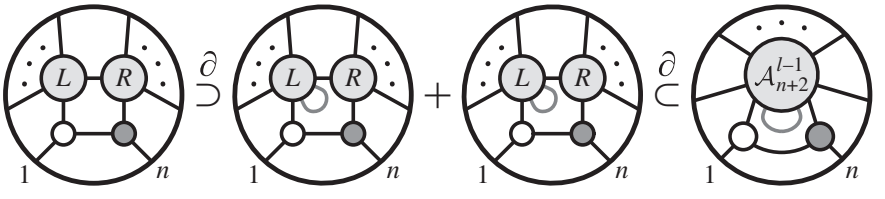
and forward-limit boundaries of the forward-limit term, respectively:



The first of these is needed by 'forward-limit' term in the BCFW recursion ansatz, and the second term is needed to complete the 'bridge' term of the recursion ansatz; to see this more clearly, notice that the second term can be redrawn more suggestively:



And so, we have shown that the induction hypothesis ensures that all the necessary forward-limit terms are generated in the boundary of the BCFW recursion formula. But as with the factorization channels studied earlier, we must show that no 'spurious' forward-limit terms are generated. Such spurious forward-limit terms can be generated by the 'bridge' term in the recursion—when the identified legs appear either between $(\widehat{1}I)$ on the left, or between $(I\widehat{n})$ on the right—or from the factorization channels of the 'forward-limit' term; these are always generated in pairs, and cancel accordingly; for example,

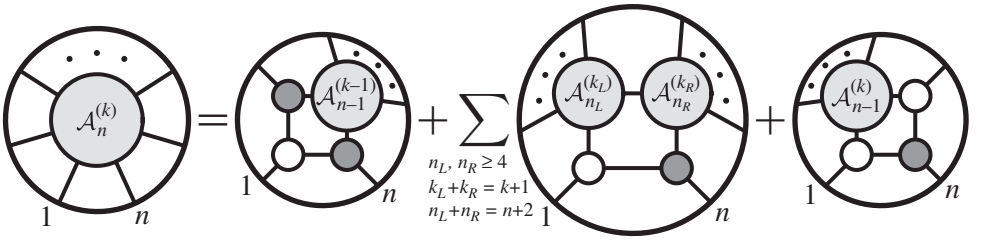


17.2 The structure of (tree) amplitudes in the Grassmannian

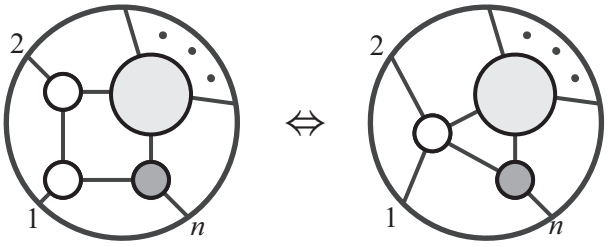
The BCFW recursion relations provide us with a powerful description of scattering amplitudes to all loop-orders. Although the tree-level recursion relations have been largely understood for nearly a decade (see e.g. [11, 74, 75, 157, 159]), its

extension to all-loop integrands remains relatively novel—and until now, has only been understood in terms of momentum-twistor variables (as described in [14]). Because of this novelty, it is worthwhile to explore some of the features of the recursion and the structures that emerge. In this subsection, we will mostly review aspects of tree-amplitudes that are well known to most practitioners; this will provide us with the background necessary to discuss some of the novelties that arise at loop-level in section 17.3.

When restricted to tree level, the recursion relations (17.2) become



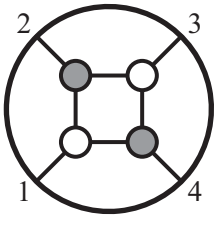
Here, we have separated the terms in the recursion that involve a 3-particle amplitude on either side of the bridge; this is because one of the 3-particle amplitudes when bridged on either side will lead to an on-shell form with vanishing support for generic kinematical data—for example, bridging $\mathcal{A}_3^{(1)}$ on the left would give



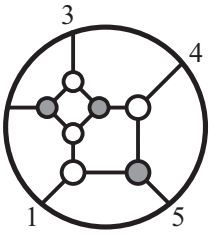
which is only non-vanishing if $\lambda_1 \propto \lambda_2$. (Moreover, it turns out that these graphs are always reducible, and so have less than the necessary $(2n-4)$ independent degrees of freedom required to solve the kinematical constraints.)

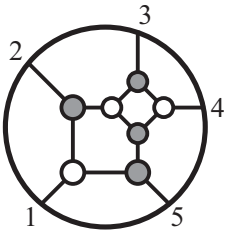
Let us begin to build intuition about the structure that arises from the recursion by considering the simplest examples. Recall that the 4-particle amplitude is entirely given by the single on-shell graph, (2.48)—the familiar 'box'

$$\mathcal{A}_4^{(2)} = \mathcal{A}_3^{(2)} \otimes \mathcal{A}_3^{(1)} =$$

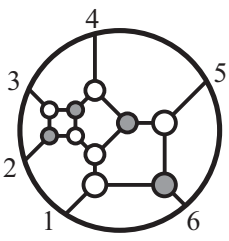


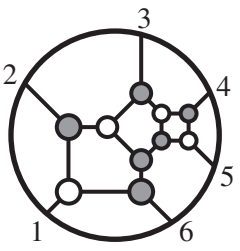
This of course follows trivially from the recursion relations. But it is not the only amplitude that is so simple: for example, the two 5-particle amplitudes are simply

$\mathcal{A}_5^{(2)} = \mathcal{A}_4^{(2)} \otimes \mathcal{A}_3^{(1)} =$ 

, $\mathcal{A}_5^{(3)} = \mathcal{A}_3^{(2)} \otimes \mathcal{A}_4^{(2)} =$ 

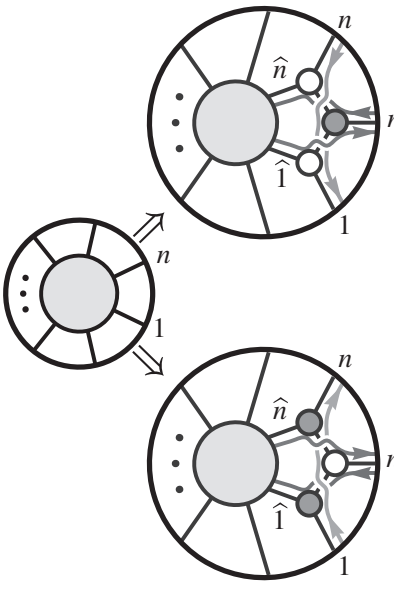
This trend continues for all MHV and $\overline{\text{MHV}}$ amplitudes, $\mathcal{A}_n^{(2)}$ and $\mathcal{A}_n^{(n-2)}$, respectively. For six particles, these amplitudes are:

$\mathcal{A}_6^{(2)} = \mathcal{A}_5^{(2)} \otimes \mathcal{A}_3^{(1)} =$ 

, $\mathcal{A}_6^{(4)} = \mathcal{A}_3^{(2)} \otimes \mathcal{A}_5^{(3)} =$ 

Thus, the BCFW-recursion *directly* represents all MHV (and $\overline{\text{MHV}}$) amplitudes as single terms—directly giving the famous formula guessed by Parke and Taylor, (8.4).

While these examples are fairly trivial, notice that in obtaining these formulae, it is natural to view the act of attaching a 3-particle amplitude across the BCFW-bridge as an operation that 'adds a particle.' This operation is of course well defined not just for the amplitude, but for any on-shell graph; thus, we have a way to add a particle in a way that 'preserves k,' $(\bullet \otimes \mathcal{A}_3^{(1)}) : G(k,n) \mapsto G(k,n+1)$, and in a way that 'increases k,' $(\mathcal{A}_3^{(2)} \otimes \bullet) : G(k,n) \mapsto G(k+1,n+1)$. These are called 'inverse-soft factors.' As a reference, these operations correspond to



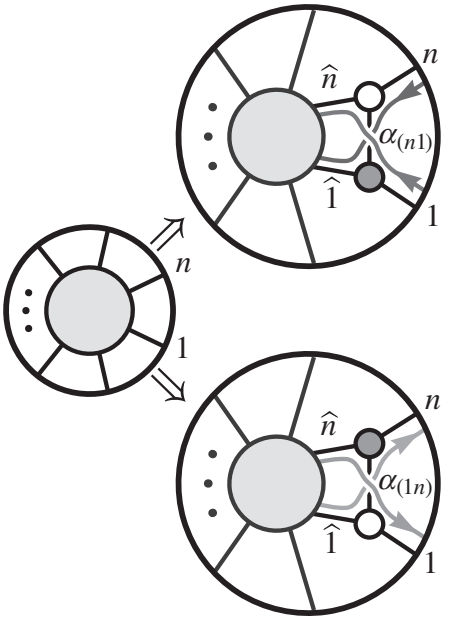
k-preserving or *holomorphic* inverse-soft factor

momentum-space	momentum-twistors
$\lambda_{\widehat{n}} = \lambda_n$ $\widetilde{\lambda}_{\widehat{n}} = \widetilde{\lambda}_n - \alpha_{(n\,n+1)}\widetilde{\lambda}_{n+1}$	$z_{\widehat{n}} = z_n$
$\lambda_{\widehat{1}} = \lambda_1$ $\widetilde{\lambda}_{\widehat{1}} = \widetilde{\lambda}_1 - \alpha_{(1\,n+1)}\widetilde{\lambda}_{n+1}$	$z_{\widehat{1}} = z_1$
$f(\cdots, n, n{+}1, 1, \cdots)$ $\Rightarrow f(\cdots, \widehat{n}, \widehat{1}, \cdots) \times$ $\delta^2\big(\lambda_{n+1} - \alpha_{(n\,n+1)}\lambda_n - \alpha_{(1\,n+1)}\lambda_1\big)$	$f(\cdots, n, n{+}1, 1, \cdots)$ $\Rightarrow f(\cdots, n, 1, \cdots)$

k-increasing or *anti-holomorphic* inverse-soft factor

momentum-space	momentum-twistors
$\lambda_{\widehat{n}} = \lambda_n + \alpha_{(n+1\,n)}\lambda_{n+1}$ $\widetilde{\lambda}_{\widehat{n}} = \widetilde{\lambda}_n$	$z_{\widehat{n}} = z_n + \alpha_{(n+1\,n)} z_{n-1}$
$\lambda_{\widehat{1}} = \lambda_1 + \alpha_{(n+1\,1)}\lambda_{n+1}$ $\widetilde{\lambda}_{\widehat{1}} = \widetilde{\lambda}_1$	$z_{\widehat{1}} = z_1 + \alpha_{(n+1\,1)} z_2$
$f(\cdots, n, n{+}1, 1, \cdots)$ $\Rightarrow f(\cdots, \widehat{n}, \widehat{1}, \cdots) \times$ $\delta^2\big(\widetilde{\lambda}_{n+1} + \alpha_{(n\,n+1)}\widetilde{\lambda}_n + \alpha_{(1\,n+1)}\widetilde{\lambda}_1\big)$	$f(\cdots, n, n{+}1, 1, \cdots)$ $\Rightarrow f(\cdots, \widehat{n}, \widehat{1}, \cdots)$ $\times [n{-}1\; n\; n{+}1\; 1\; 2]$

(Here, the $\widetilde{\eta}$'s transform identically to the $\widetilde{\lambda}$'s.) Each of these can be seen to follow from the action of two successive BCFW-bridges:



BCFW-bridge '$(n\,1)$'

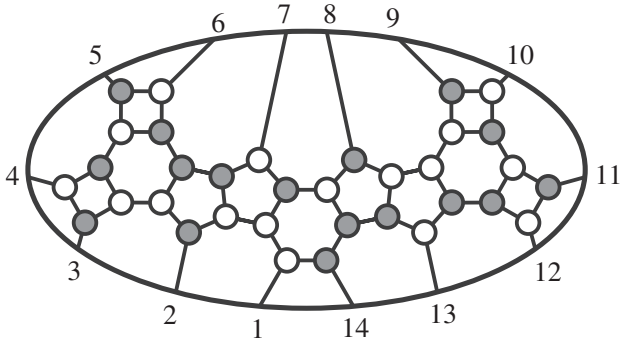
momentum-space	momentum-twistors
$\lambda_{\widehat{n}} = \lambda_n$ $\widetilde{\lambda}_{\widehat{n}} = \widetilde{\lambda}_n - \alpha_{(n1)}\widetilde{\lambda}_1$	$z_{\widehat{n}} = z_n$
$\lambda_{\widehat{1}} = \lambda_1 + \alpha_{(n1)}\lambda_n$ $\widetilde{\lambda}_{\widehat{1}} = \widetilde{\lambda}_1$	$z_{\widehat{1}} = z_1 + \alpha_{(n1)} z_2$

BCFW-bridge '$(1\,n)$'

momentum-space	momentum-twistors
$\lambda_{\widehat{n}} = \lambda_n + \alpha_{(1n)}\lambda_1$ $\widetilde{\lambda}_{\widehat{n}} = \widetilde{\lambda}_n$	$z_{\widehat{n}} = z_n + \alpha_{(1n)} z_{n-1}$
$\lambda_{\widehat{1}} = \lambda_1$ $\widetilde{\lambda}_{\widehat{1}} = \widetilde{\lambda}_1 - \alpha_{(1n)}\widetilde{\lambda}_n$	$z_{\widehat{1}} = z_1$

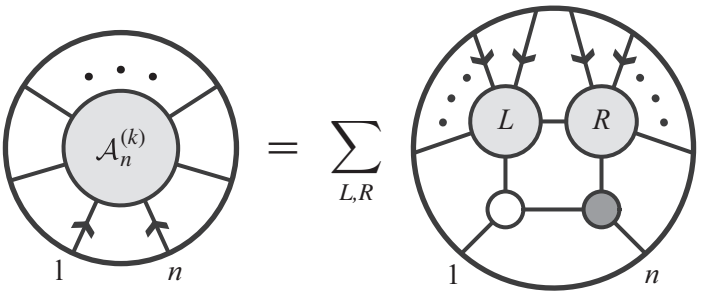
Notice that whenever an on-shell graph has a leg a such that $\sigma(a-1) = a+1$ or $\sigma(a+1) = a-1$ we can view it as having been obtained by adding particle a to a lower-point graph using a k-preserving or k-increasing inverse-soft factor, respectively. In such cases, a is said to be an 'inverse-soft factor'; and any on-shell graph which can be constructed by successively adding particles to a 3-particle amplitude using inverse-soft factors is said to be 'inverse-soft *constructible*'.

The notion of inverse-soft constructibility proves useful because the auxiliary variables associated with any inverse-soft factor can be completely fixed by the associated δ-function constraint, making it very easy to recursively eliminate all the auxiliary, Grassmannian degrees of freedom. It turns out that for 13 or fewer legs, *all* on-shell forms generated by the tree-level recursion relations—*regardless* of how lower-point amplitudes are themselves recursed—are inverse-soft constructible. (Actually, this is only true if a single BCFW parity is used for recursion; using both parities (white-to-black and black-to-white), non inverse-soft constructible terms can occur starting with 10 legs.) However, for 14 or more particles, some objects can be generated by the recursion relations that are *not* inverse-soft constructible, such as the following possible contribution to the 14-particle N^{3}MHV tree-amplitude (labeled by {4,7,6,10,16,17,14,15,12,13,19,23,22,25}):



$$((\mathcal{A}_3^{(2)}\otimes(\mathcal{A}_4^{(2)}\otimes\mathcal{A}_4^{(2)}))\otimes\mathcal{A}_3^{(1)})\otimes(\mathcal{A}_3^{(2)}\otimes((\mathcal{A}_4^{(2)}\otimes\mathcal{A}_4^{(2)})\otimes\mathcal{A}_3^{(1)})) \tag{17.3}$$

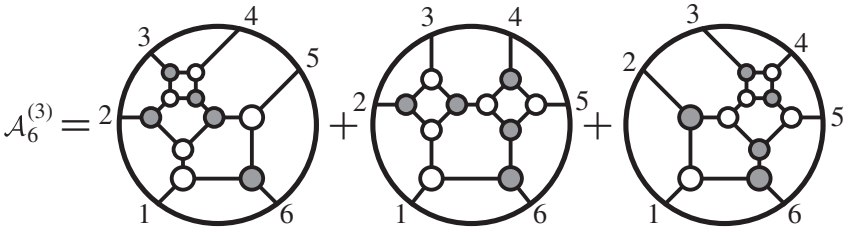
Notice that this graph was generated by always using *internal edges* to recurse the objects appearing across the BCFW-bridge—$(\widehat{1}I)$ on the left and $(I\widehat{n})$ on the right. (We should mention in passing that if one *always* recurses the lower-point amplitudes according to the marked legs as follows,



(17.4)

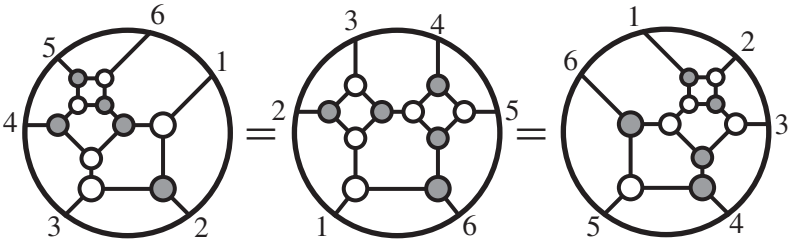
then *all* tree-amplitudes will be given in terms of only inverse-soft constructible graphs. This corresponds to the recursion 'scheme' $\{-2,2,0\}$ of reference [160].)

As described in Chapter 11, the first amplitude given as the combination of several on-shell graphs is $\mathcal{A}_6^{(3)}$, the six-particle NMHV tree-amplitude. This is given by three terms, $\mathcal{A}_5^{(3)}\otimes\mathcal{A}_3^{(1)}$, $\mathcal{A}_4^{(2)}\otimes\mathcal{A}_4^{(2)}$, and $\mathcal{A}_3^{(2)}\otimes\mathcal{A}_5^{(2)}$:



{4, 5, 6, 8, 7, 9} {3, 5, 6, 7, 8, 10} {4, 6, 5, 7, 8, 9} (17.5)

Although the on-shell graphs of each contribution appear quite different, it is easy to see from the permutations that they are all cyclically related to one another



{3, 5, 6, 7, 8, 10} {3, 5, 6, 7, 8, 10} {3, 5, 6, 7, 8, 10} (17.6)

The on-shell differential form drawn above—labeled by $\{3,5,6,7,8,10\}$—was given directly in terms of the kinematical variables $\lambda, \widetilde{\lambda}$ in equation (8.7). Because each term is cyclically related, if we use 'r' to denote the operation that 'rotates' all particle labels forward by 1, we can write the entire tree amplitude as follows:

$$\mathcal{A}_6^{(3)} = (1 + r^2 + r^4) \frac{\delta^{3\times 4}(C^* \cdot \widetilde{\eta})\, \delta^{2\times 2}\big(\lambda \cdot \widetilde{\lambda}\big)}{\langle 23\rangle[56](\langle 34\rangle[64] + \langle 53\rangle[56]) s_{456} (\langle 61\rangle[64] + \langle 15\rangle[45])\langle 12\rangle[45]},$$

where the matrix C^* was given in (8.6).

Although the precise set of on-shell graphs obtained using the BCFW recursion relations can vary considerably depending on which legs of the lower-point amplitudes are used for *their* recursion, the number of terms is of course scheme-independent. It is a relatively simple exercise to show that the number of BCFW terms in the tree-amplitude $\mathcal{A}_n^{(k)}$ is

$$\frac{1}{n-3}\binom{n-3}{k-1}\binom{n-3}{k-2}. \tag{17.7}$$

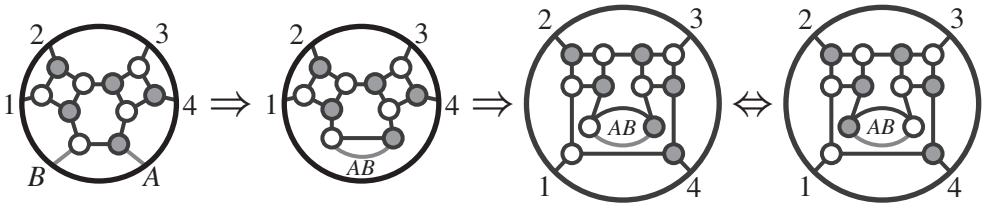
17.3 Canonical coordinates for loop integrands

The all-loop generalization of the BCFW recursion relations was first described in [14] where it was formulated in terms of explicit operations acting directly on the 'functions' of momentum-twistor variables obtained after eliminating the auxiliary Grassmannian degrees of freedom. This led to formulae for the 'loop integrands' in the form of a 'standard' loop-integration measure $d^4\ell$ weighted by some rational function of the loop-momentum ℓ, or equivalently a function one the space of lines AB in momentum-twistor space with measure $d^4z_A d^4z_B/GL(2)$. When viewed as rational functions in this way, much of the underlying structure is hidden. However, by viewing each loop-momentum's degrees of freedom as arising from canonical coordinates in the auxiliary Grassmannian, the integration measure is automatically generated in a much more illuminating, 'canonical' form:

as a wedge-product of "$d\log$" factors. The fact that loop amplitude integrands can be written in such a form—a fact which is essentially obvious using canonical coordinates on the Grassmannian—is far from obvious from any other method to compute scattering amplitudes.

We will postpone a systematic discussion of the loop amplitude integrands generated by the recursion relations (17.2) until a future work. Here, we merely want to demonstrate its most important physical implications through the context of simple examples. We first describe how one-loop integrands are generated by the recursion, using the case of MHV for illustration. At the end of this subsection, we will briefly describe the features observed for higher-loop amplitudes.

Let us begin with the simplest of all one-loop amplitudes, the 4-particle MHV amplitude. As there are no 3-particle one-loop integrands to appear in the 'bridge' term of the recursion, the 4-particle one-loop integrand is entirely generated as the forward-limit of the 6-particle NMHV tree-amplitude, $\mathcal{A}_6^{(3)}$. Let us denote the two particles identified in the forward-limit by $(A\,B)$, and use these two legs as the pair singled out in the recursion of the 6-particle tree. Of the three terms appearing in the tree-amplitude $\mathcal{A}_6^{(3)}$, (17.5), only one is non-vanishing in the forward-limit (a fact that we will demonstrate momentarily); the forward-limit of the $\mathcal{A}_4^{(2)}\otimes\mathcal{A}_4^{(2)}$-term is

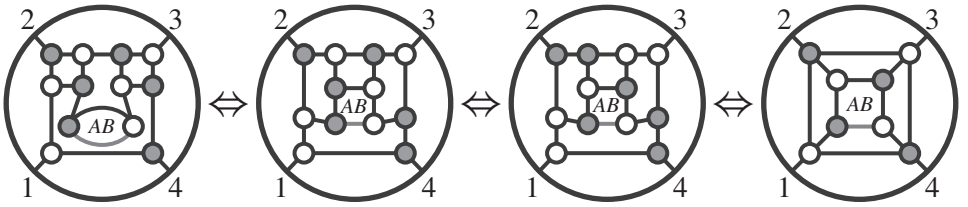


(the last move in this sequence was made only to make subsequent transformations more transparent). It is easy to see that this diagram has four faces beyond that of the simple box, and thus four extra integration variables. Using reduction, we can of course reduce this diagram to the box, giving us the integrand. We can relate this new form of the integrand to a more familiar form, by identifying the usual loop momentum "ℓ" as

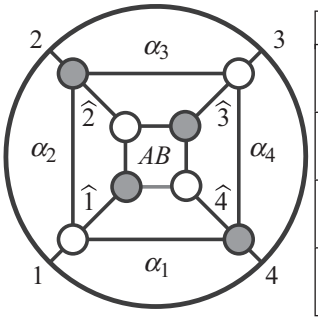
$$\ell = \alpha\lambda_1\widetilde{\lambda}_4 + \lambda_{AB}\widetilde{\lambda}_{AB}, \tag{17.8}$$

where $\lambda_{AB}\widetilde{\lambda}_{AB}$ is the momentum of the highlighted line in the figure above. We can of course determine λ_{AB} and $\widetilde{\lambda}_{AB}$ in terms of the variables associated with the graph, and in this way trade the four 'extra' variables for those that parameterize ℓ.

While this is a straightforward exercise, it is more illuminating to carry out the reduction in a different way. We can use moves to give the on-shell diagram a different representation—as a sequence of BCFW-bridges on a core 4-particle amplitude:



(in the last transformation, several mergers were made). This allows us to think of the object as the usual box, but with 'BCFW-shifted kinematical data,' given by



bridge	BCFW shift
(14)	$\widetilde{\lambda}_4 \mapsto \widetilde{\lambda}_4 + \alpha_1\widetilde{\lambda}_1$ $\widetilde{\lambda}_1 \mapsto \widetilde{\lambda}_1 - \alpha_1\widetilde{\lambda}_4$
(12)	$\widetilde{\lambda}_2 \mapsto \widetilde{\lambda}_2 + \alpha_2\widetilde{\lambda}_1$ $\widetilde{\lambda}_1 \mapsto \widetilde{\lambda}_1 - \alpha_2\widetilde{\lambda}_2$
(32)	$\widetilde{\lambda}_2 \mapsto \widetilde{\lambda}_2 + \alpha_3\widetilde{\lambda}_3$ $\widetilde{\lambda}_3 \mapsto \widetilde{\lambda}_3 - \alpha_3\widetilde{\lambda}_2$
(34)	$\widetilde{\lambda}_4 \mapsto \widetilde{\lambda}_4 + \alpha_4\widetilde{\lambda}_3$ $\widetilde{\lambda}_3 \mapsto \widetilde{\lambda}_3 - \alpha_4\widetilde{\lambda}_4$

$$\Rightarrow \begin{cases} \widehat{1} \begin{cases} \lambda_{\widehat{1}} = \lambda_1 \\ \widetilde{\lambda}_{\widehat{1}} = \widetilde{\lambda}_1 - \alpha_1\widetilde{\lambda}_4 - \alpha_2\widetilde{\lambda}_2 \end{cases} \\ \widehat{2} \begin{cases} \lambda_{\widehat{2}} = \lambda_2 + \alpha_2\lambda_1 + \alpha_3\lambda_3 \\ \widetilde{\lambda}_{\widehat{2}} = \widetilde{\lambda}_2 \end{cases} \\ \widehat{3} \begin{cases} \lambda_{\widehat{3}} = \lambda_3 \\ \widetilde{\lambda}_{\widehat{3}} = \widetilde{\lambda}_3 - \alpha_3\widetilde{\lambda}_2 - \alpha_4\widetilde{\lambda}_4 \end{cases} \\ \widehat{4} \begin{cases} \lambda_{\widehat{4}} = \lambda_4 + \alpha_1\lambda_1 + \alpha_4\lambda_3 \\ \widetilde{\lambda}_{\widehat{4}} = \widetilde{\lambda}_4 \end{cases} \end{cases}$$

Thus, the integrand is simply $d\log(\alpha_1) \wedge \cdots \wedge d\log(\alpha_4)$ times the shifted four-particle amplitude,

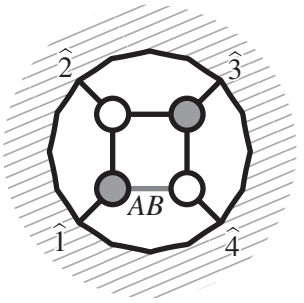
$$\frac{d\alpha_1}{\alpha_1}\frac{d\alpha_2}{\alpha_2}\frac{d\alpha_3}{\alpha_3}\frac{d\alpha_4}{\alpha_4}\frac{\delta^{2\times 4}\big(\lambda\cdot\widetilde{\eta}\big)}{\langle\widehat{1}\widehat{2}\rangle\langle\widehat{2}\widehat{3}\rangle\langle\widehat{3}\widehat{4}\rangle\langle\widehat{4}\widehat{1}\rangle}\delta^{2\times 2}\big(\lambda\cdot\widetilde{\lambda}\big). \tag{17.9}$$

If we strip off the (unshifted) Parke–Taylor prefactor, the integrand for the one-loop ratio function—the loop amplitude divided by the tree-amplitude—is simply [161]

$$\begin{aligned}
\frac{\mathcal{A}_4^{(2),1}}{\mathcal{A}_4^{(2),0}} &= \frac{d\alpha_1}{\alpha_1}\frac{d\alpha_2}{\alpha_2}\frac{d\alpha_3}{\alpha_3}\frac{d\alpha_4}{\alpha_4}\frac{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle}{\langle\widehat{1}\widehat{2}\rangle\langle\widehat{2}\widehat{3}\rangle\langle\widehat{3}\widehat{4}\rangle\langle\widehat{4}\widehat{1}\rangle} \\
&= \frac{d\alpha_1}{\alpha_1}\frac{d\alpha_2}{\alpha_2}\frac{d\alpha_3}{\alpha_3}\frac{d\alpha_4}{\alpha_4}\frac{\langle 12\rangle}{\langle 12\rangle + \alpha_3\langle 13\rangle}\frac{\langle 23\rangle}{\langle 23\rangle + \alpha_2\langle 13\rangle} \\
&\quad \times \frac{\langle 34\rangle}{\langle 34\rangle + \alpha_1\langle 31\rangle}\frac{\langle 41\rangle}{\langle 41\rangle + \alpha_4\langle 31\rangle} \\
&= d\log\left(\frac{\alpha_1\langle 34\rangle}{\langle 34\rangle + \alpha_1\langle 31\rangle}\right) d\log\left(\frac{\alpha_2\langle 23\rangle}{\langle 23\rangle + \alpha_2\langle 13\rangle}\right) \\
&\quad \times d\log\left(\frac{\alpha_3\langle 12\rangle}{\langle 12\rangle + \alpha_3\langle 13\rangle}\right) d\log\left(\frac{\alpha_4\langle 41\rangle}{\langle 41\rangle + \alpha_4\langle 31\rangle}\right)
\end{aligned}$$

which is manifestly in a '$d$log'-form.

Now, we can determine λ_{AB} and $\widetilde{\lambda}_{AB}$ very simply in terms of the bridge variables. For a general box, the internal momentum on the bridge can be given in terms of the external data according to



(17.10)

which allows us to identify

$$\lambda_{AB}\widetilde{\lambda}_{AB} = \frac{\widehat{\langle 12\rangle}}{\widehat{\langle 42\rangle}}\lambda_{\widehat{4}}\widetilde{\lambda}_{\widehat{1}}. \tag{17.11}$$

And so in summary, the relation between the BCFW-bridge variables α_i and the usual loop-momentum variables is given by

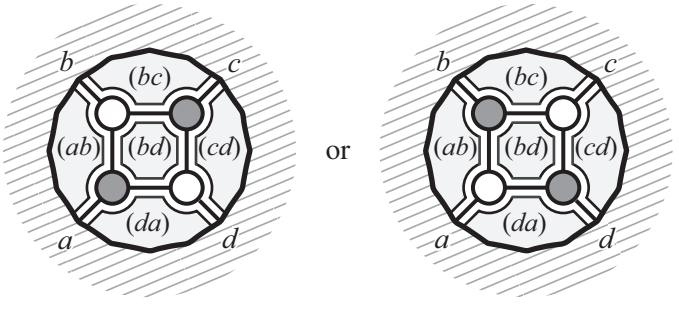
$$\ell = \frac{\widehat{\langle 12\rangle}}{\widehat{\langle 42\rangle}}\lambda_{\widehat{4}}\widetilde{\lambda}_{\widehat{1}} + \alpha_1\lambda_1\widetilde{\lambda}_4. \tag{17.12}$$

Using this change of variables, it is straightforward to recast the integrand (17.10) in the form we gave earlier in Chapter 2:

$$d\log\left(\frac{\ell^2}{(\ell-\ell^*)^2}\right)d\log\left(\frac{(\ell+p_1)^2}{(\ell-\ell^*)^2}\right)d\log\left(\frac{(\ell+p_1+p_2)^2}{(\ell-\ell^*)^2}\right)d\log\left(\frac{(\ell-p_4)^2}{(\ell-\ell^*)^2}\right), \tag{17.13}$$

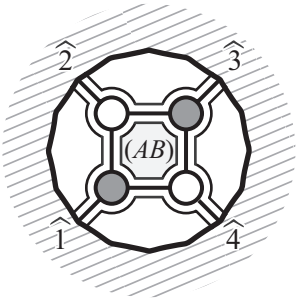
where $\ell^* = \frac{\langle 12\rangle}{\langle 42\rangle}\lambda_4\widetilde{\lambda}_1$.

We can also interpret exactly the same pictures in momentum-twistor space. Recall that a BCFW-bridge '$(a-1\,a)$'—a white-to-black vertex from $a-1$ to a—has the effect of shifting the momentum twistor $z_a \mapsto z_a + \alpha z_{a+1}$ where α is the bridge variable. Generally speaking, lines in momentum-twistor space are associated with the faces of the momentum-space on-shell graph; we will not review these ideas here, but let us briefly summarize that the regions of a four-point box are associated with the lines in momentum-twistor space as indicated below:



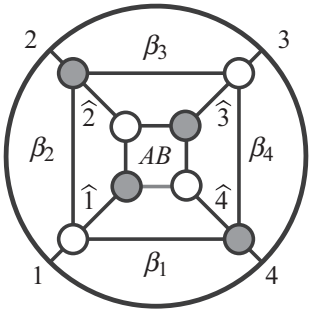
(17.14)

Now, this means that if we identify the four unfixed degrees of freedom with the line (AB) in momentum-twistor space, we see that it corresponds to the line $\widehat{(24)}$ in



(17.15)

Performing the same sequence of shifts as before, but now using momentum-twistor variables, we find:



bridge	BCFW shift
(14)	$z_4 \mapsto z_4 + \beta_1 z_3$
(12)	$z_2 \mapsto z_2 + \beta_2 z_3$
(32)	$z_2 \mapsto z_2 + \beta_3 z_1$
(34)	$z_4 \mapsto z_4 + \beta_4 z_1$

$$\Rightarrow \left\{ \begin{array}{l} z_{\widehat{1}} = z_1 \\ z_{\widehat{2}} = z_2 + \beta_2 z_3 + \beta_3 z_1 \\ z_{\widehat{3}} = z_3 \\ z_{\widehat{4}} = z_4 + \beta_1 z_3 + \beta_4 z_1 \end{array} \right\}$$

This uniquely fixes the auxiliary, Grassmannian parameters β_i in terms of the momentum-twistor line (AB) according to:

$$\beta_1 = \frac{\langle AB41\rangle}{\langle AB13\rangle},\ \beta_2 = \frac{\langle AB12\rangle}{\langle AB31\rangle},\ \beta_3 = \frac{\langle AB23\rangle}{\langle AB31\rangle},\ \text{and}\ \ \beta_4 = \frac{\langle AB34\rangle}{\langle AB13\rangle}. \tag{17.16}$$

With this identification, we can rewrite the integrand in terms of the four auxiliary variables in momentum-twistor space as

$$\begin{aligned} &d\log(\beta_1)\cdots d\log(\beta_4) \\ &\quad = d\log\left(\frac{\langle AB41\rangle}{\langle AB13\rangle}\right) d\log\left(\frac{\langle AB12\rangle}{\langle AB31\rangle}\right) d\log\left(\frac{\langle AB23\rangle}{\langle AB31\rangle}\right) d\log\left(\frac{\langle AB34\rangle}{\langle AB13\rangle}\right). \end{aligned} \tag{17.17}$$

If we recast this expression as an integration measure on the space of lines (AB) in momentum-twistor space, we find that

$$\begin{aligned} &d\log\left(\frac{\langle AB41\rangle}{\langle AB13\rangle}\right) d\log\left(\frac{\langle AB12\rangle}{\langle AB31\rangle}\right) d\log\left(\frac{\langle AB23\rangle}{\langle AB31\rangle}\right) d\log\left(\frac{\langle AB34\rangle}{\langle AB13\rangle}\right) \\ &\quad = \frac{\langle d^2 z_A AB\rangle \langle d^2 z_B AB\rangle \langle 1234\rangle \langle 2341\rangle}{\langle AB12\rangle \langle AB23\rangle \langle AB34\rangle \langle AB41\rangle}, \end{aligned} \tag{17.18}$$

which is precisely the familiar form of the integrand given in reference [14].

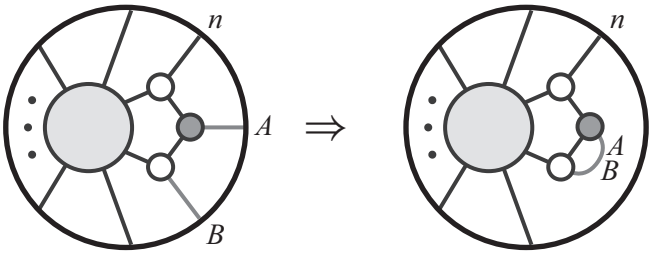
Before moving on to the case of the n-particle MHV one-loop integrand, let us go back and understand why only one of the three terms in the six-particle

NMHV tree amplitude survived the forward-limit, as the reason will prove quite instructive. Let us choose to always represent the $(n+2)$-point tree-amplitude appearing in the forward-limit using the BCFW recursion that deforms legs (AB). Recall that the tree-amplitude recursion can be broken into three parts as in (17.3):

$$\begin{aligned} &1.\ \text{a } k\text{-preserving inverse-soft factor:} \quad \mathcal{A}^{(k)}_{n-1}\otimes\mathcal{A}^{(1)}_{3};\\ &2.\ \text{a } k\text{-increasing inverse-soft factor:} \quad \mathcal{A}^{(2)}_{3}\otimes\mathcal{A}^{(k-1)}_{n-1};\\ \text{and}\quad &3.\ \text{terms for which } n_L, n_R \geq 4. \end{aligned} \tag{17.19}$$

Of these, it is not hard to see that if (AB) are the distinguished legs of the bridge, the first two contributions listed above *always* vanish. More precisely, any on-shell form for which A *or* B is an inverse-soft factor will vanish in the forward-limit. (We should notice that 'A or B being an inverse-soft factor' is a *sufficient* condition for an on-shell form to vanish in the forward limit, but not a *necessary* one.)

Let us now see why any contributions to the lower-loop amplitude where A or B is an inverse-soft factor will vanish. Consider the forward-limit of a term for which A is a k-preserving inverse-soft factor (the argument is the same in all other cases):



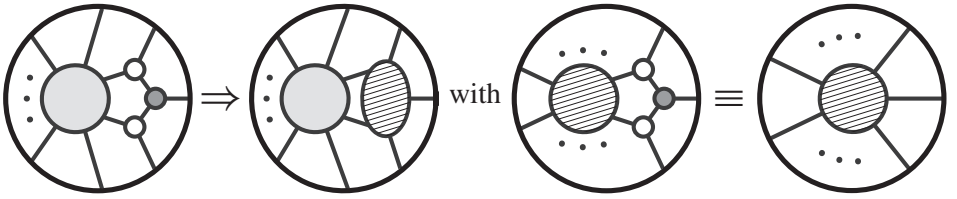
(17.20)

Notice that the kinematical constraints associated with the middle—black—vertex requires that λ_A be expandable in terms of λ_B and λ_n; but in the forward-limit, we identify λ_A with λ_B, which implies that $\lambda_{AB} \propto \lambda_n$. As such, the kinematical constraints do not allow for there to be any unfixed degrees of freedom associated with λ_{AB} (which should represent loop-integration degrees of freedom).

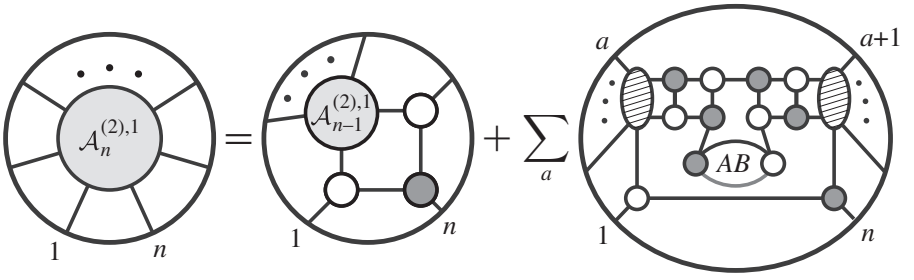
We are now prepared to determine the n-point MHV one-loop integrand in general. The bridge-term always contributes a term $\mathcal{A}^{(2),1}_{n-1}\otimes\mathcal{A}^{(1)}_{3}$, which is simply a *k-preserving* inverse-soft factor adding n to the $(n-1)$-point one-loop amplitude; more interesting are the forward-limit terms. These come from the forward-limit of $\mathcal{A}^{(3),0}_{n+2}$; among the terms that contribute to the higher-point NMHV tree-amplitude, we have seen that only those obtained from bridging $\mathcal{A}^{(2)}_{n_L}\otimes\mathcal{A}^{(2)}_{n_R}$ with $n_L, n_R \geq 4$ contribute.

Because k-preserving inverse-soft factors act trivially in momentum-twistors, and the left- and right-amplitudes appearing in the NMHV tree-amplitudes are trivially chains of inverse-soft factors, it will be useful to define the notion of an "MHV region" obtained by any number of successive k-preserving inverse-soft

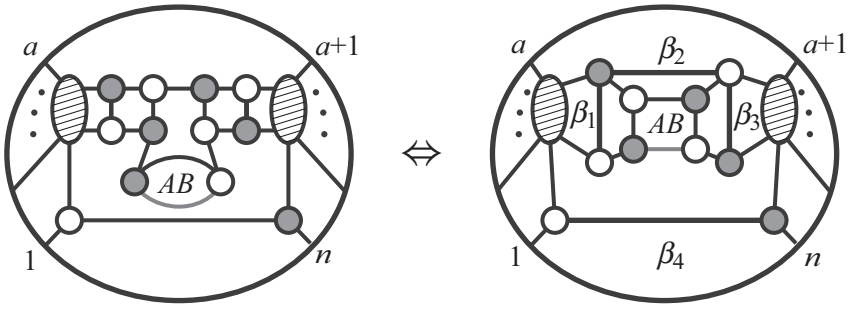
factors:



Allowing for such MHV regions in our diagrammatic expansion, we see that the one-loop MHV integrand is given by



We can rearrange the NMHV forward-limit contributions as we did above in order to make manifest the sequence of BCFW-bridges that parameterize the extra degrees of freedom:



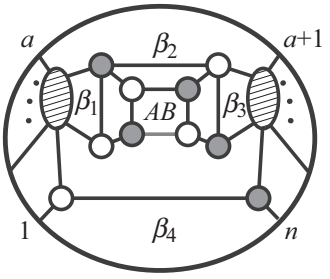
Proceeding as before, we can identify the shifted momentum-twistors appearing in the box as

$$(AB) = (\widehat{a\,n}) \quad \text{with} \quad \left\{\begin{matrix} z_{\widehat{a}} = z_a + \beta_1 z_{a+1} + \beta_2 z_1 \\ z_{\widehat{n}} = z_n + \beta_3 z_1 + \beta_4 z_{n-1} \end{matrix}\right\}, \tag{17.21}$$

which allows us to recast the BCFW-bridge variables β_i in terms of the line (AB):

$$\beta_1 = \frac{\langle AB\,1\,a\rangle}{\langle AB\,a+1\,1\rangle},\ \beta_2 = \frac{\langle AB\,a\,a+1\rangle}{\langle AB\,a+1\,1\rangle},\ \beta_3 = \frac{\langle AB\,n-1\,n\rangle}{\langle AB\,1\,n-1\rangle},\ \text{and}\ \beta_4 = \frac{\langle AB\,n\,1\rangle}{\langle AB\,1\,n{-}1\rangle}.$$

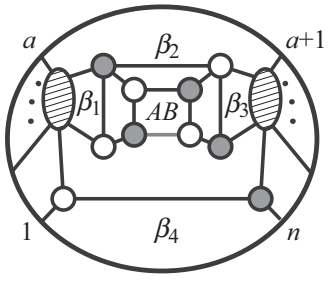
Therefore, we see that the forward-limit terms are given by



$$= d\log(\beta_1)d\log(\beta_2)d\log(\beta_3)d\log(\beta_4)$$

$$= d\log\left(\frac{\langle AB\,1a\rangle}{\langle AB\,1a+1\rangle}\right) d\log\left(\frac{\langle AB\,aa+1\rangle}{\langle AB\,1a+1\rangle}\right) d\log\left(\frac{\langle AB\,n-1n\rangle}{\langle AB\,1n-1\rangle}\right) d\log\left(\frac{\langle AB\,n1\rangle}{\langle AB\,1n-1\rangle}\right).$$

Quite amazingly, if we recast this integration measure directly in terms of the line (AB), we see that this is equivalent to



$$= \frac{\langle d^2z_A\,AB\rangle\langle d^2z_B\,AB\rangle\langle AB\,(1\,a\,a+1)\bigcap(1\,n{-}1\,n)\rangle^2}{\langle AB\,1a\rangle\langle AB\,aa+1\rangle\langle AB\,a+11\rangle\langle AB\,1n{-}1\rangle\langle AB\,n{-}1n\rangle\langle AB\,n1\rangle}$$
$$\equiv K[a;n-1].$$

We have obtained this result entirely by manipulating pictures of on-shell diagrams; of course the result precisely matches the form obtained by direct computation, using the methods of [14], where all MHV one-loop integrands were given in the form

$$\mathcal{A}_n^{(2),1} = \sum_{1<a<b<n} K[a;b]. \tag{17.22}$$

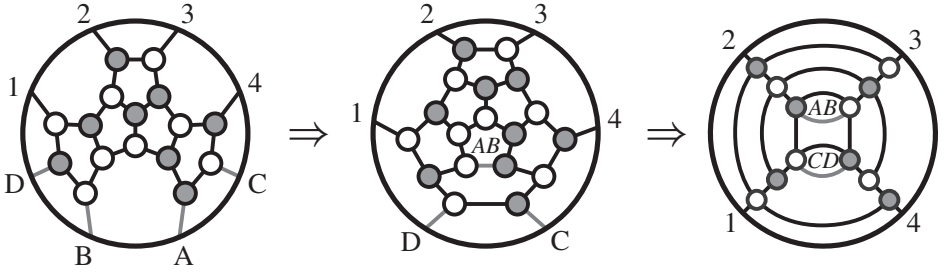
Before moving on to multi-loop integrands, it is worth mentioning that for one-loop integrands, so long as the forward-limits are taken of tree-amplitudes obtained by BCFW-deforming the identified legs $(A\,B)$, it turns out that the obvious k-preserving and k-increasing inverse-soft factors are the *only* terms that vanish in the forward limit; this allows us to conclude that

$$\#\text{ BCFW terms in the one-loop amplitude } \mathcal{A}_n^{(k),1} \text{ is : } \binom{n-2}{k}\binom{n-2}{k-2}. \tag{17.23}$$

(It turns out that this counting holds regardless of how the forward-limit terms are recursed—even though it is generally not known how to identify beforehand which terms will vanish if (AB) are not singled out for the recursion. Beyond 1-loop, however, the number of non-vanishing contributions is not invariant, and depends sensitively on how the lower-loop amplitudes are recursed.)

When expressing tree-amplitudes and their forward-limits in terms of canonical coordinates on the auxiliary Grassmannian, it is obvious that all loop integrands can be—and are most naturally—expressed in such a '*d*log'-representation. Although in principle we have all the tools necessary to construct such formulae for all amplitudes—and although the BCFW recursion relations (17.2) is *dramatically* more efficient that any representation obtained using 'traditional' methods (such as Feynman diagrams)—even the simplest 2-loop integrands would require more space to write completely than would be warranted for the purpose of illustration.

Let us therefore content ourselves with considering one simple example of a contribution to the 4-particle 2-loop integrand that arises as the double forward-limit of the contributions to the 8-particle N^{2}MHV tree-amplitude, that of $\left(\mathcal{A}_4^{(2)}\otimes\mathcal{A}_4^{(2)}\right)\otimes\mathcal{A}_4^{(2)}$:



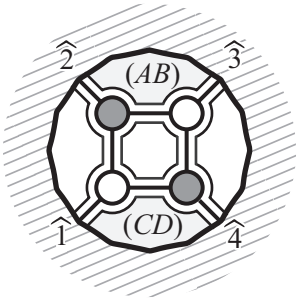
In the last step, we have made liberal use of square moves and merge/unmerge operations to bring the tree amplitude into a form that exposes a sequence of recognizable BCFW-bridges, which themselves encode the additional degrees of freedom.

Using the tools described in [14] to compute this contribution directly as a 'function' of lines (AB) and (CD) in momentum-twistor space, the following rational integrand is found:

$$\frac{\langle d^2z_A\,AB\rangle\langle d^2z_B\,AB\rangle\langle d^2z_C\,CD\rangle\langle d^2z_D\,CD\rangle\langle 1234\rangle^3\langle AB(CD)\bigcap(341)1\rangle^2}{\langle AB14\rangle\langle AB1(123)\cap(CD)\rangle\langle AB1(234)\cap(CD)\rangle\langle ABY1\rangle\langle ABCD\rangle\langle CD34\rangle},$$

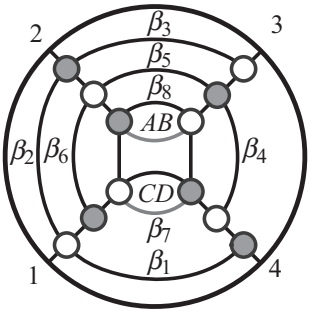
where we defined the twistor $z_Y\equiv(((CD(341)\cap(AB))\cap(12))34)\cap(CD)$. While the expression above is of course obtained in a straightforward way, it is obviously rather complicated and not particularly illuminating. Moreover, as written in the form given above—as a rational integrand—it is not at all obvious that there exists *any* change of variables for which it becomes simply the wedge-product of eight logarithmic factors. But from our present perspective, the existence of such a change of coordinates is an obvious consequence of the Grassmannian formulation of the initial tree-amplitude; and it will be instructive to see how this remarkable connection is realized.

To be extremely concrete, we want to identify the lines (AB) and (CD) as parameterizing the region-momenta according to



with
$$\begin{aligned}(AB) &= (\widehat{2}\widehat{3})\\ (CD) &= (\widehat{4}\widehat{1})\end{aligned} \tag{17.24}$$

We can find the shifted momentum-twistors $z_{\widehat{a}}$ by performing the successive BCFW-shifts that are obvious from the way the double forward-limit graph is drawn:



bridge	BCFW shift
(14)	$z_4 \mapsto z_4 + \beta_1 z_3$
(12)	$z_2 \mapsto z_2 + \beta_2 z_3$
(32)	$z_2 \mapsto z_2 + \beta_3 z_1$
(43)	$z_3 \mapsto z_3 + \beta_4 z_2$
(23)	$z_3 \mapsto z_3 + \beta_5 z_4$
(21)	$z_1 \mapsto z_1 + \beta_6 z_4$
(14)	$z_4 \mapsto z_4 + \beta_7 z_3$
(32)	$z_2 \mapsto z_2 + \beta_8 z_1$

$$\Longrightarrow \left\{ \begin{aligned} z_{\widehat{1}} &= z_1 + \beta_5(z_4 + \beta_1 z_3)\\ z_{\widehat{2}} &= z_2 + \beta_2 z_3 + \beta_3 z_1\\ &\quad + \beta_8(z_1 + \beta_5(z_4 + \beta_1 z_3))\\ z_{\widehat{3}} &= z_3 + \beta_4(z_2 + \beta_2 z_3 + \beta_3 z_1)\\ &\quad + \beta_6(z_4 + \beta_1 z_3)\\ z_{\widehat{4}} &= z_4 + \beta_1 z_3 + \beta_7\beta_6(z_4 + \beta_1 z_3)\\ &\quad + \beta_7(z_3 + \beta_4(z_2 + \beta_2 z_3 + \beta_3 z_1)) \end{aligned} \right\}$$

One can readily verify that quite remarkably, with this change of variables, the complicated expression given above becomes simply

$$d\log(\beta_1) \wedge \cdots \wedge d\log(\beta_8). \tag{17.25}$$

17.4 The transcendentality of loop amplitudes

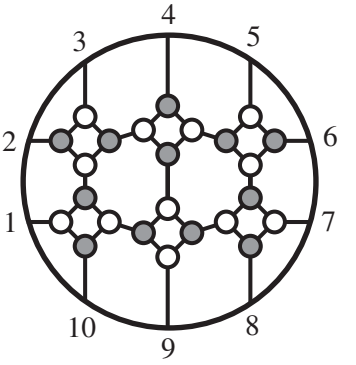
The integrand obtained from the BCFW recursion relations allows us to draw some important general conclusions about the structure of the final, integrated expressions for the amplitude. Let us start with MHV amplitudes. As we have seen, all the BCFW terms at l loops can be written in the form

$$\mathcal{A}_n^{(2),l} = \mathcal{A}_n^{(2),0} \times \prod_{i=1}^{4l} d\log(\beta_i). \tag{17.26}$$

The first and most obvious point to observe is that the integrand has *only* logarithmic singularities! There are *no* "*sub-leading*" pieces of the integrand with less than the maximal number of logarithmic singularities. At 1-loop, this (together with dual conformal invariance) tells us the famous fact that the loop amplitude only depends on "box" integrals, and doesn't involve any triangles, bubbles, or rational pieces [162, 163].

As we have stressed a number of times, the fact that the integrand has only logarithmic singularities is not at all obvious from inspection of the actual rational functions involved in sufficiently high loop-amplitude integrands, where there don't seem to be enough "obvious" singularities in cutting propagators, and so singularities must emerge as "composites". By contrast, the positive Grassmannian story makes this fact completely obvious. Intuitively, this guarantees that after integration, the l-loop MHV amplitudes can always be expressed as a sum of polylogarithms of transcendentality $2l$. The reason is roughly that discontinuities of the amplitude are related to unitarity cuts that put pairs of particles on-shell; thereby computing partial residues of the integrand. Taking $2l$ discontinuities gives the leading singularity "1", which has no further discontinuities. These amplitudes are thus "pure"—not polluted by lower-transcendentality terms, which would arise from pieces of the integrand without purely logarithmic singularities. This has long been conjectured to be true for MHV amplitudes in connection to the maximal transcendentality principle of [164]. We see that the property needed of the integrand to guarantee this is a trivial consequence of the $d$log form.

Beyond MHV amplitudes, we know that the integrated amplitudes can involve more complicated functions than polylogarithms. For instance, as pointed out in reference [92], the two-loop, 10-point N^3MHV amplitude includes a contribution from a function whose sevenfold discontinuity is proportional to the following on-shell form:



$$\{7,6,10,9,8,12,11,15,14,13\} \tag{17.27}$$

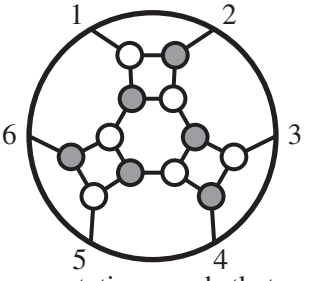
This on-shell graph corresponds to a 17-dimensional cell in $G(5,10)$; the kinematical constraints will fix this to an integral over one degree of freedom (interpreted as the 'hepta-cut' of the two-loop integrand). The component amplitude involving

$$(\tilde{\eta}_1^1\tilde{\eta}_1^2)(\tilde{\eta}_2^1\tilde{\eta}_2^2)(\tilde{\eta}_3^2\tilde{\eta}_3^3)(\tilde{\eta}_4^2\tilde{\eta}_4^3)(\tilde{\eta}_5^2\tilde{\eta}_5^3)(\tilde{\eta}_6^3\tilde{\eta}_6^4)(\tilde{\eta}_7^3\tilde{\eta}_7^4)(\tilde{\eta}_8^4\tilde{\eta}_8^1)(\tilde{\eta}_9^4\tilde{\eta}_9^1)(\tilde{\eta}_{10}^4\tilde{\eta}_{10}^1)$$

vanishes exactly at tree level and one-loop, but is non-vanishing at two-loops, vanishing on all the positroid cells in the boundary of (17.27). Moreover, *all* 16-dimensional cells in $G(5,10)$ vanish for this component. Because of this, the only contour integral available must enclose the Jacobian resulting from the kinematical constraints; this Jacobian generically involves the square root of an

irreducible quartic, implying that (at least for this component) the seven-fold discontinuity of the 2-loop integrand is an elliptic integral.

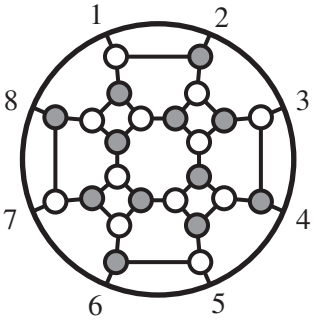
We can understand the difference between MHV and higher-k amplitudes from the Grassmannian. Recall that cells of dimensionality $(2n-4)$ are fully localized by the kinematical constraints. Since for MHV amplitudes, $\dim(G(2,n)) = (2n-4)$, *all* of the unfixed degrees of freedom associated with 'loop-momenta' are associated with faces that can always be removed by reduction (as there no irreducible graphs with more faces than that of the top-cell). Beyond MHV, however, the reduction of on-shell diagrams can result in cells of higher dimensionality than $(2n-4)$. For example, consider the top-cell of $G(3,6)$:



(17.28)

Here, we have chosen a representative graph that makes it clear that it can be associated with a triple-cut of the 6-particle amplitude at 1-loop. The $9-8=1$ degree of freedom of the top-cell not fixed by the kinematical constants can always be interpreted as the single integration variable of a triple-cut integral.

Similarly, the top-cell of $G(4,8)$ is 16-dimensional, while the kinematical constraints can be used to isolate only 12 degrees of freedom; therefore, the top-cell on $G(4,8)$ can be viewed as an on-shell differential form with *four* unfixed degrees of freedom—which can in fact be interpreted as the four degrees of freedom of a 'loop-integrand' at 1-loop. Indeed, we can represent the top-cell by



(17.29)

Therefore beyond MHV, while the integration measures are purely $d$log's, some free integration variables are *inside* the Grassmannian, and must be localized by the kinematic constraints. This is the reason why more complicated functions can appear after integration. However, it is clear that for fixed n and k, the functions can't get arbitrarily more complicated at high loop orders. The reason is that at most $\dim(G(k,n)) - (2n-4)$ of the integration variables can remain 'entangled' in the Grassmannian (meaning that they cannot be pulled off as overall $d$log factors

in the measure via bubble-reduction); at arbitrarily high loop order, all but a finite number of these auxiliary degrees of freedom must be associated with the more trivial factors in the measure arising from bubble-reduction.

Actually, it is easy to see that, for NMHV amplitudes, the integrations that are "stuck" in the Grassmannian can easily be removed, preserving the $d\log$ form, and thus that all NMHV amplitudes are also polylogarithms. Let us illustrate with the top-cell of $G(3,6)$; it is convenient to work in momentum-twistor language, where this maps to the top-cell of $G(1,6)$. On the support of the (ordinary) δ-functions, we have a 1-form that we can represent as

$$[1\,2\,3\,4\,(5+\beta 6)]d\log(\beta), \tag{17.30}$$

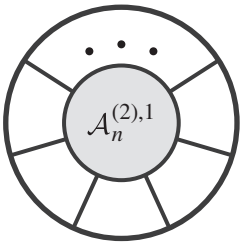
However, we can use the identity among the 5-brackets, (11.3), to rewrite this as

$$\begin{aligned}[1\,2\,3\,4\,5]d\log(\beta)+[2\,3\,4\,5\,6]d\log\left(\beta-\frac{\langle 2345\rangle}{\langle 2346\rangle}\right)+[3\,4\,5\,6\,1]d\log\left(\beta-\frac{\langle 3451\rangle}{\langle 3461\rangle}\right)\\ +[4\,5\,6\,1\,2]d\log\left(\beta-\frac{\langle 4512\rangle}{\langle 4612\rangle}\right)+[5\,6\,1\,2\,3]d\log\left(\beta-\frac{\langle 5123\rangle}{\langle 6123\rangle}\right).\end{aligned} \tag{17.31}$$

In this way, we have removed the integration variable from inside the Grassmannian and decomposed the result into a sum of terms, each of which is in canonical form. The same thing can be done for the top-cell of any NMHV amplitude, since the "internal" variables always occur linearly. Things can start becoming nontrivial at N^2MHV, where square roots first make an appearance, and as we've seen concretely above, already for 10-particle N^3MHV amplitudes, elliptic integrals do make an appearance.

The on-shell BCFW-representation of loop-integrands delivers them manifestly in a canonical, $d\log$-form; but having noted that the integrand can be put in this form, it is natural to wonder if this is a consequence of the BCFW-representation, or a more general result. For instance, in reference [165], extremely compact, *local* forms of many integrands were found; can these also be written in terms of integrands with only logarithmic singularities? The answer is "yes": the $d\log$ form is a *general* property of "pure" integrands with unit leading singularities. Let us briefly demonstrate this fact with two examples: local forms of the MHV 1- and 2-loop integrands.

In [165], the 1-loop MHV integrand was given in the local form



$$= \sum_{a<b<a} I_X[a;b], \tag{17.32}$$

where $I_X[a;b]$ denotes the integrand

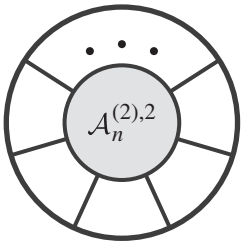
$$I_X[a;b] \equiv \frac{\langle ABd^2z_A\rangle\langle ABd^2z_B\rangle\ \langle AB(a-1aa+1)\bigcap(b-1bb+1)\rangle\langle Xab\rangle}{\langle AB\,a-1\,a\rangle\langle AB\,a\,a+1\rangle\langle AB\,b-1\,b\rangle\langle AB\,b\,b+1\rangle\langle AB\,X\rangle}, \quad (17.33)$$

and where X is an arbitrary reference line in momentum-twistor space (spanned by any pair of twistors). Remarkably, it turns out that $I_X[a;b]$ *can* be expressed in canonical form:

$$d\log\left(\frac{\langle ABa-1a\rangle}{\langle AB\,X\rangle}\right)d\log\left(\frac{\langle ABaa+1\rangle}{\langle AB\,X\rangle}\right)$$
$$\times\, d\log\left(\frac{\langle ABb-1b\rangle}{\langle AB(aX)\bigcap(b-1bb+1)\rangle}\right)d\log\left(\frac{\langle ABbb+1\rangle}{\langle AB(aX)\bigcap(b-1bb+1)\rangle}\right). \quad (17.34)$$

(Note that while this form appears to break the symmetry between a and b, the form is of course symmetrical.)

Similarly, it was found in [165] that two-loop MHV integrand could be written in the following local form:



$$= \sum_{a<b<c<d<a} I[a,b;c,d] \quad (17.35)$$

where $I[a,b;c,d]$ denotes the integrand proportional to,

$$\frac{\langle AB(a-1aa+1)\bigcap(b-1bb+1)\rangle\langle CD(c-1cc+1)\bigcap(d-1dd+1)\rangle}{\langle AB\,a-1a\rangle\langle AB\,aa+1\rangle\langle AB\,b-1b\rangle\langle AB\,bb+1\rangle\langle AB\,CD\rangle\langle CD\,c-1c\rangle\langle CD\,cc+1\rangle\langle CD\,d-1d\rangle\langle CD\,dd+1\rangle}. \quad (17.36)$$

But it turns out that this integrand can also be written in canonical form:

$$d\log\left(\frac{\langle AB\,a-1a\rangle}{\langle AB\,CD\rangle}\right)d\log\left(\frac{\langle AB\,aa+1\rangle}{\langle AB\,CD\rangle}\right)d\log\left(\frac{\langle AB\,b-1b\rangle}{\langle AB(aCD)\bigcap(b-1bb+1)\rangle}\right)$$
$$\times\, d\log\left(\frac{\langle AB\,bb+1\rangle}{\langle AB(aCD)\bigcap(b-1bb+1)\rangle}\right)d\log\left(\frac{\langle CD\,c-1c\rangle}{\langle CD\,ab\rangle}\right)d\log\left(\frac{\langle CD\,cc+1\rangle}{\langle CD\,ab\rangle}\right)$$
$$\times\, d\log\left(\frac{\langle CD\,d-1d\rangle}{\langle CD(abc)\bigcap(d-1dd+1)\rangle}\right)d\log\left(\frac{\langle CD\,dd+1\rangle}{\langle CD(abc)\bigcap(d-1dd+1)\rangle}\right). \quad (17.37)$$

18

Outlook

We have explored much of the remarkable physics and mathematics of scattering amplitudes in planar $\mathcal{N} = 4$, as seen through the lens of on-shell diagrams as the primary objects of study. Let us conclude by making some brief comments on further avenues of research.

One immediate extension of our work is the continued study of theories with $\mathcal{N} < 4$ supersymmetry, whose most basic features we sketched out in Chapter 15. For $\mathcal{N} \geq 1$, we expect that all-loop BCFW recursion holds just as for $\mathcal{N} = 4$, together with its realization in terms of on-shell diagrams. For non-supersymmetric theories, the forward-limit of tree amplitudes are singular, and thus don't directly give us the single-cuts of the loop-integrand [81]. More thought is needed to establish a connection between on-shell diagrams and the full amplitude, though it is likely that fully understanding the on-shell diagrams will continue to play an important role in determining $\mathcal{N} = 0$ amplitudes as well.

The general connection between on-shell diagrams and the Grassmannian has nothing to do with any particular theory, only with the general picture of amalgamating basic three-particle amplitudes, and the connection to the positive Grassmannian in particular holds for any planar theory of massless particles in 4 dimensions. Only the form on the Grassmannian changes from theory to theory. As briefly discussed in Chapter 15, the essential physical novelty of gauge theories with $\mathcal{N} \leq 2$ supersymmetry is the presence of UV divergences. The most physical, Wilsonian way to think about UV divergences makes critical use of off-shell ideas, and so a major challenge is finding the correct way of thinking about such physics in a directly on-shell language. It is fascinating to see that the UV and IR singularities, together with UV/IR decoupling, are reflected directly in on-shell diagrams through simple structures in the Grassmannian. A clear goal would be to understand the physics of the renormalization group along these lines.

Another obvious extension is to push beyond the planar limit, starting already with $\mathcal{N} = 4$; in this case, there is no longer an obvious notion of "the loop integrand," and thus we must learn how to establish a connection between

on-shell diagrams and the full scattering amplitude along the lines of the BCFW construction in the planar limit. It is also very likely that on-shell ideas can be used to determine other observables in gauge theories beyond scattering amplitudes, including all correlation functions. These objects also have discontinuities and cuts, and the on-shell diagrams for leading singularities of form-factors and correlation functions are exactly the same as the (in general nonplanar) on-shell diagrams we have been considering. The structure of cuts has already proved powerful in determining form-factors [166]. For scattering amplitudes, we have seen that off-shell notions like virtual loop integration variables can be fully understood in on-shell terms. It is tempting to try and compute completely off-shell objects like correlation functions in the same way.

Moving further afield, as alluded to in Chapter 16, the basic mathematical structures we have encountered in scattering amplitudes have also recently made an appearance in apparently completely different physical settings, related to conformal blocks for higher Toda theories [36, 46], wall-crossing [47, 48], various versions of the AGT conjecture [167], scattering amplitudes at strong coupling [168], and soliton solutions of the KP equation [55–57]. The identical graphical structure has also appeared in the construction of $\mathcal{N} = 1$ super-conformal field theories associated with quiver gauge theories (see e.g. [53,54]). The combinatorial classification of on-shell diagrams and these planar $\mathcal{N} = 1$ super-conformal field theories coincide perfectly. It would be interesting to see if the rest of the structure we have been seeing—especially the connection with the positive Grassmannian—has a natural interpretation as well.

There is a unifying theme running through the physics and mathematics we have been discussing. We have an object—the positive Grassmannian—which is fundamentally defined by global properties, either as a real space, by demanding all ordered minors to be positive, or as a complex space, by specifying linear dependencies between consecutive vectors. However, quite remarkably, the best way of building up these objects (albeit in a highly redundant way) is through the amalgamation of elementary building blocks.

For scattering amplitudes, amalgamation representations have a direct physical interpretation as on-shell diagrams. For $\mathcal{N} = 1$ gauge theories, they correspond to gluing together gauge groups with bi-fundamental content to generate more complicated quiver gauge theories. For scattering amplitudes, it is physically clear why we should be interested in complicated on-shell diagrams, since they are ultimately needed to compute the amplitude to all-loop order. But what is physically important about complicated quiver gauge theories? One possible answer is that precisely these sort of quiver gauge theories, with an infinite number of sites and links, occur in the deconstruction of the still mysterious $(2,0)$ and little string theories in six dimensions [169]. It would be fascinating to use the powerful new machinery for studying these quivers to try and learn more about the dynamics

of the underlying six-dimensional theories, which would perhaps shed some light on a more direct physical reason for the appearance of the same Grassmannian structure in seemingly vastly different settings.

We have seen that scattering amplitudes in $(1+1)$, $(2+1)$, and $(3+1)$ dimensions are described by various interpretations of permutations and associated structures in the Grassmannian. It is natural to ask whether other variations of these mathematical ideas might have a physical interpretation. There is one natural further specialization of the positive Grassmannian we have not discussed, which in fact goes back to the historical roots of the subject: the study of totally positive matrices. Here, one considers $(n \times n)$ square matrices M with positive determinant, and studies the space where all its $(m \times m)$-minors are non-negative. This classical problem was studied by Gantmacher and Krein [43] and Schoenberg [44] in the 1930, where the stratification was found to be determined by *pairs* of permutations σ_1 and σ_2. This theory is a special case of the positive Grassmannian $G(n,2n)$. Consider cells where the first n columns of the $(n \times 2n)$ C matrix are linearly independent, and also the second n columns are linearly independent. We can then gauge-fix C to the form

$$C = (1_{n\times n} \,|\, M_{n\times n}), \tag{18.1}$$

where M is a positive matrix. Let us label the first n columns $1,\ldots,n$ and the second n columns by $1',\ldots,n'$. It is clear that, for example, $\sigma(1) = a'$ for some a' in the second set of columns, since 1 can not be in the span of $\{2,\ldots,n\}$, given that the first n columns are linearly independent. This is true for all the other columns in the first set—that is, $\sigma(a) = b'$. Similarly, $\sigma(a') = b$. Thus, we see that our permutation naturally breaks into two pieces, mapping $(1,\ldots,n) \mapsto (1',\ldots,n')$ and vice versa. It would be nice to find a physical interpretation for the subclass of on-shell diagrams associated with these pairs of permutations.

We have also seen that a reliable guide to the Grassmannian structure associated with scattering amplitudes is to find a Grassmannian interpretation of the space of external kinematical data. In four dimensions, the λ- and $\widetilde{\lambda}$-planes are represented by points in $G(2,n)$. In three dimensions, the λ-plane is an element of the null orthogonal Grassmannian. What happens in higher dimensions? The description of the external kinematical data in six dimensions is particularly simple [170]. The complexified Lorentz group can be taken to be $SO(5,1) \sim SL(4)$, and a null momentum vector can be represented as an antisymmetric (4×4) tensor p^{IJ} of vanishing determinant. The complexified little group is $SL(2) \times SL(2)$. As such, we can express the momentum p_a^{IJ} of particle a as

$$p_a^{IJ} = \epsilon^{\alpha\beta}\lambda^I_{a\alpha}\lambda^J_{a\beta}\big(= \epsilon^{\dot\alpha\dot\beta}\widetilde{\lambda}^I_{a\dot\alpha}\widetilde{\lambda}^J_{a\dot\beta}\big). \tag{18.2}$$

Note the similarity to ordinary spinor-helicity variables—except that here, the $\alpha, \dot\alpha$ indices aren't Lorentz indices as familiar from four dimensions, but are instead

indices of the $SL(2) \times SL(2)$ *little group*. We can group all the λ's for the particles $a = 1, \ldots, n$ together into a $(4 \times 2n)$ matrix,

$$\Lambda_A^I = \left(\lambda_{11}^I \, \lambda_{12}^I \, \lambda_{21}^I \, \lambda_{22}^I \cdots \lambda_{n1}^I \, \lambda_{n2}^I\right). \tag{18.3}$$

Momentum conservation is then the statement that

$$\Lambda_A^I \Lambda_B^J J^{AB} = 0 \quad \text{where} \quad J \equiv \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}. \tag{18.4}$$

Thus, in close parallel with $(2+1)$ dimensions, the external data in six dimensions are associated with a point in the null symplectic Grassmannian [171]. It would be interesting to see if this structure has any role to play in six-dimensional physics.

Let us close by returning to a number of concrete open directions of research flowing more directly from the ideas presented here.

In this book, we have given a complete classification and understanding of all reduced on-shell diagrams, whose invariant content is captured by the permutation associated with the left-right paths. Amongst other things, all terms occurring in the tree-level BCFW recursion relations are reduced graphs, and indeed, the recursion can be described purely combinatorially as a simple and canonical "bridging" of permutations. We have however also seen that non-reduced on-shell diagrams are also physically important, directly giving the loop integrand. Of course, the non-reduced graphs for the loop integrand arise from merging adjacent legs of higher-point reduced graphs, which we understand completely. Nonetheless, it would clearly be interesting and important to try and extend the classification of the on-shell diagrams to non-reduced graphs as well; in other words, we would like to understand all the invariants on non-reduced graphs that can be related by merges and square moves. Obviously the left-right path permutations are still invariants, but there are clearly further invariants as well. For instance, suppose we have two non-reduced graphs with exactly the same permutation, but where the first graph has a bad double-crossing between the two paths starting at a and b, while the second has a bad double-crossing for a different pair of paths starting at c and d. Clearly square moves and merges can't connect these two diagrams. It is plausible that the complete set of invariants involves the permutation together with other labels characterizing the pattern of intersections of the left-right paths. Finding a complete classification will be very important, not least because it would allow us to cast the BCFW construction of all-loop integrands in completely combinatorial terms.

We have seen that the all-loop integrand is naturally presented in a "$d\log$" form. This form begs to be integrated, indeed (most naïvely of course), these forms integrate to zero! The integrals don't vanish because of branch cuts in the

arguments of the logarithms, on the real contour of integration. This leads to novel ways of performing the loop integrations directly in space-time, which will be pursued in future work.

Finally, the BCFW construction of scattering amplitudes in the Grassmannian still leaves something to be desired. It is not entirely satisfying to give the scattering amplitude a fundamentally recursive definition. Put another way, we have yet to see locality and unitarity fully emerge from more primitive principles in a completely satisfactory way. We would like to have a direct definition of scattering amplitudes, linked to the Grassmannian, making all the symmetries manifest, and discover their singularities in the form of factorization and forward-limits as an emergent property. This is bound to be linked to the "polytope picture" studied in [128, 172]. This line of thought will certainly be taken up again with our vastly improved understanding of the positive Grassmannian in hand.

References

[1] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. B. Goncharov, A. Postnikov, *et al.*, "Scattering Amplitudes and the Positive Grassmannian," arXiv:1212.5605 [hep-th].

[2] S. Weinberg, *The Quantum Theory of Fields. Vol. 1: Foundations*. Cambridge University Press, 1995.

[3] N. Seiberg and E. Witten, "Electric-Magnetic Duality, Monopole Condensation, and Confinement in $\mathcal{N}=2$ Supersymmetric Yang–Mills Theory," *Nucl. Phys.* **B426** (1994) 19–52, arXiv:hep-th/9407087.

[4] N. Seiberg, "Electric-Magnetic Duality in Supersymmetric non-Abelian Gauge Theories," *Nucl. Phys.* **B435** (1995) 129–146, arXiv:hep-th/9411149.

[5] A. Kapustin and E. Witten, "Electric-Magnetic Duality And The Geometric Langlands Program," *Commun. Num. Theor. Phys.* **1** (2007) 1–236, arXiv:hep-th/0604151.

[6] J. M. Maldacena, "The Large-*N* Limit of Superconformal Field Theories and Supergravity," *Adv. Theor. Math. Phys.* **2** (1998) 231–252, arXiv:hep-th/9711200.

[7] M. L. Mangano and S. J. Parke, "Multiparton Amplitudes in Gauge Theories," *Phys. Rept.* **200** (1991) 301–367, arXiv:hep-th/0509223.

[8] L. J. Dixon, "Calculating Scattering Amplitudes Efficiently," arXiv:hep-ph/9601359.

[9] F. Cachazo and P. Svrcek, "Lectures on Twistor Strings and Perturbative Yang–Mills Theory," *PoS* **RTN2005** (2005) 004, arXiv:hep-th/0504194.

[10] Z. Bern, L. J. Dixon, and D. A. Kosower, "On-Shell Methods in Perturbative QCD," *Annals Phys.* **322** (2007) 1587–1634, arXiv:0704.2798 [hep-ph].

[11] B. Feng and M. Luo, "An Introduction to On-Shell Recursion Relations," *Front. Phys.* **5, 7** (2011) 533–575, arXiv:1111.5759 [hep-th].

[12] R. Roiban, M. Spradlin, and A. Volovich, "Scattering Amplitudes in Gauge Theories: Progress and Outlook," *J. Phys.* **A44** (2011) no. 45, 1.

[13] N. Beisert, C. Ahn, L. F. Alday, Z. Bajnok, J. M. Drummond, *et al.*, "Review of AdS/CFT Integrability: An Overview," *Lett. Math. Phys.* **99** (2012) 3–32, arXiv:1012.3982 [hep-th].

[14] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, S. Caron-Huot, and J. Trnka, "The All-Loop Integrand For Scattering Amplitudes in Planar $\mathcal{N}=4$ SYM," *JHEP* **1101** (2011) 041, arXiv:1008.2958 [hep-th].

[15] N. Arkani-Hamed, F. Cachazo, C. Cheung, and J. Kaplan, "A Duality For The *S*-Matrix," *JHEP* **1003** (2010) 020, arXiv:0907.5418 [hep-th].

[16] L. Mason and D. Skinner, "Dual Superconformal Invariance, Momentum Twistors and Grassmannians," *JHEP* **0911** (2009) 045, arXiv:0909.0250 [hep-th].

[17] N. Arkani-Hamed, F. Cachazo, and C. Cheung, "The Grassmannian Origin Of Dual Superconformal Invariance," *JHEP* **1003** (2010) 036, arXiv:0909.0483 [hep-th].

[18] J. L. Bourjaily, J. Trnka, A. Volovich, and C. Wen, "The Grassmannian and the Twistor String: Connecting All Trees in $\mathcal{N}=4$ SYM," *JHEP* **1101** (2011) 038, arXiv:1006.1899 [hep-th].
[19] J. Drummond, "Review of AdS/CFT Integrability, Chapter V.2: Dual Superconformal Symmetry," *Lett. Math. Phys.* **99** (2012) 481–505, arXiv:1012.4002 [hep-th].
[20] A. B. Goncharov, "A Simple Construction of Grassmannian Polylogarithms," *Adv. Math.* (April, 2013), arXiv:0908.2238 [math.AG].
[21] A. B. Goncharov, M. Spradlin, C. Vergu, and A. Volovich, "Classical Polylogarithms for Amplitudes and Wilson Loops," *Phys. Rev. Lett.* **105** (2010) 151605, arXiv:1006.5703 [hep-th].
[22] S. Caron-Huot, "Superconformal Symmetry and Two-Loop Amplitudes in Planar $\mathcal{N}=4$ Super Yang–Mills," *JHEP* **1112** (2011) 066, arXiv:1105.5606 [hep-th].
[23] L. F. Alday, "Some Analytic Results for Two-Loop Scattering Amplitudes," *JHEP* **1107** (2011) 080, arXiv:1009.1110 [hep-th].
[24] L. J. Dixon, J. M. Drummond, and J. M. Henn, "Bootstrapping the Three-Loop Hexagon," *JHEP* **1111** (2011) 023, arXiv:1108.4461 [hep-th].
[25] P. Heslop and V. V. Khoze, "Wilson Loops @ 3-Loops in Special Kinematics," *JHEP* **1111** (2011) 152, arXiv:1109.0058 [hep-th].
[26] L. J. Dixon, J. M. Drummond, and J. M. Henn, "Analytic Result for the Two-Loop Six-Point NMHV Amplitude in $\mathcal{N}=4$ Super Yang–Mills Theory," *JHEP* **1201** (2012) 024, arXiv:1111.1704 [hep-th].
[27] J. L. Bourjaily, A. DiRe, A. Shaikh, M. Spradlin, and A. Volovich, "The Soft-Collinear Bootstrap: $\mathcal{N}=4$ Yang–Mills Amplitudes at Six and Seven Loops," *JHEP* **1203** (2012) 032, arXiv:1112.6432 [hep-th].
[28] B. Eden, P. Heslop, G. P. Korchemsky, and E. Sokatchev, "Constructing the Correlation Function of Four Stress-Tensor Multiplets and the Four-Particle Amplitude in $\mathcal{N}=4$ SYM," *Nucl. Phys.* **B862** (2012) 450–503, arXiv:1201.5329 [hep-th].
[29] S. Caron-Huot and S. He, "Jumpstarting the All-Loop *S*-Matrix of Planar $\mathcal{N}=4$ Super Yang–Mills," *JHEP* **1207** (2012) 174, arXiv:1112.1060 [hep-th].
[30] A. Sever, P. Vieira, and T. Wang, "OPE for Super Loops," *JHEP* **1111** (2011) 051, arXiv:1108.1575 [hep-th].
[31] A. Sever, P. Vieira, and T. Wang, "From Polygon Wilson Loops to Spin Chains and Back," *JHEP* **1212** (2012) 065, arXiv:1208.0841 [hep-th].
[32] N. Gromov, V. Kazakov, and P. Vieira, "Exact Spectrum of Anomalous Dimensions of Planar $\mathcal{N}=4$ Supersymmetric Yang–Mills Theory," *Phys. Rev. Lett.* **103** (2009) 131601, arXiv:0901.3753 [hep-th].
[33] G. Lusztig, "Total Positivity in Partial Flag Manifolds," *Representation Theory* **2** (1998) 70–78.
[34] G. Lusztig, "Total Positivity in Reductive Groups," in *Lie Theory and Geometry, In Honor of B. Kostant*, vol. 123 of *Prog. in Math.*, pp. 531–569. Boston: Birkhauser 1994.
[35] S. Fomin and A. Zelevinsky, "Cluster Algebras. I. Foundations," *J. Amer. Math. Soc.* **15** (2002) no. 2, 497–529, arXiv:math/0104151.
[36] V. V. Fock and A. B. Goncharov, "Moduli Spaces of Local Systems and Higher Teichmüller Theory," *Publ. Math. IHES* (2006) no. 103, 1–212, arXiv:math.AG/0311149.
[37] V. V. Fock and A. B. Goncharov, "Cluster Ensembles, Quantization and the Dilogarithm," *Ann. Sci. L'Ecole Norm. Sup.* (2009) , arXiv:math.AG/0311245.
[38] V. V. Fock and A. B. Goncharov, "Cluster $\mathcal{X}$-Varieties, Amalgamation and Poisson–Lie Groups," in *Algebraic Geometry and Number Theory, Dedicated to Drinfeld's 50$^{\text{th}}$ birthday*, pp. 27–68. Boston: Birkhauser 2006. arXiv:math.RT/0508408.
[39] A. Postnikov, "Total Positivity, Grassmannians, and Networks," arXiv:math/0609764.

[40] A. Knutson, T. Lam, and D. Speyer, "Positroid Varieties: Juggling and Geometry," *Compositio Mathematica* **149** (2013) no. 10, 1710–1752, arXiv:1111.3660 [math.AG].
[41] A. B. Goncharov and R. Kenyon, "Dimers and Cluster Integrable Systems," *Ann. Sci. L'Ecole Norm. Sup.* **46** (2013) no. 4, 747–813, arXiv:1107.5588 [math.AG].
[42] D. Thurston, "From Dominoes to Hexagons," arXiv:math/0405482.
[43] F. Gantmacher and M. Krein, "Sur les Matrices Oscillatores," *CR Acad. Sci. Paris* **201** (1935).
[44] I. J. Schoenberg, "Über Variationsvermindernde Lineare Transformationen," *Math. Z.* **32** (1930) 321–322.
[45] O. Aharony, O. Bergman, D. L. Jafferis, and J. Maldacena, "$\mathcal{N}=6$ Superconformal Chern–Simons Matter Theories, $M2$-Branes and Their Gravity Duals," *JHEP* **0810** (2008) 091, arXiv:0806.1218 [hep-th].
[46] V. V. Fock and A. B. Goncharov, "The Quantum Dilogarithm and Quantisation of Cluster Varieties," *Inventiones Math.* **175** (2009) 223–286, arXiv:math.QA/0702397.
[47] M. Kontsevich and Y. Soibelman, "Stability Structures, Motivic Donaldson-Thomas Invariants and Cluster Transformations," arXiv:0811.2435 [math.AG].
[48] D. Gaiotto, G. W. Moore, and A. Neitzke, "Wall-Crossing in Coupled $2d$-$4d$ Systems," arXiv:1103.2598 [hep-th].
[49] D. Xie, "Network, Cluster Coordinates and $\mathcal{N}=2$ Theory I," arXiv:1203.4573 [hep-th].
[50] S. Franco, "Bipartite Field Theories: from D-Brane Probes to Scattering Amplitudes," *JHEP* **1211** (2012) 141, arXiv:1207.0807 [hep-th].
[51] D. Xie and M. Yamazaki, "Network and Seiberg Duality," *JHEP* **1209** (2012) 036, arXiv:1207.0811 [hep-th].
[52] D. Xie, "Network, Cluster Coordinates and $\mathcal{N}=2$ Theory II: Irregular Singularity," arXiv:1207.6112 [hep-th].
[53] J. J. Heckman, C. Vafa, D. Xie, and M. Yamazaki, "String Theory Origin of Bipartite SCFTs," *JHEP* **1305** (2013) 148, arXiv:1211.4587 [hep-th].
[54] S. Franco, D. Galloni, and R.-K. Seong, "New Directions in Bipartite Field Theories," *JHEP* **1306** (2013) 032, arXiv:1211.5139 [hep-th].
[55] Y. Kodama and L. Williams, "KP Solitons, Total Positivity, and Cluster Algebras," *Proc. Natl. Acad. Sci. USA* **108** (2011) no. 22, 8984–8989, arXiv:1105.4170 [math.CO].
[56] Y. Kodama and L. Williams, "KP Solitons and Total Positivity for the Grassmannian," *Inventiones Mathematicae* (2014) 1–63, arXiv:1106.0023 [math.CO].
[57] Y. Kodama and L. Williams, "A Deodhar Decomposition of the Grassmannian and the Regularity of KP solitons," arXiv:1204.6446 [math.CO].
[58] V. Pestun, "Localization of Gauge Theory on a Four-Sphere and Supersymmetric Wilson Loops," *Commun. Math. Phys.* **313** (2012) 71–129, arXiv:0712.2824 [hep-th].
[59] D. Gaiotto and E. Witten, "S-Duality of Boundary Conditions in $\mathcal{N}=4$ Super Yang–Mills Theory," *Adv. Theor. Math. Phys.* **13** (2009) , arXiv:0807.3720 [hep-th].
[60] B. L. van der Waerden, "Spinoranalyse," *Nach. Ges. Wiss. Göttingen Math.-Phys.* **1** (1929) 100–109.
[61] F. A. Berends, R. Kleiss, P. De Causmaecker, R. Gastmans, and T. T. Wu, "Single Bremsstrahlung Processes in Gauge Theories," *Phys. Lett.* **B103** (1981) 124.
[62] P. De Causmaecker, R. Gastmans, W. Troost, and T. T. Wu, "Multiple Bremsstrahlung in Gauge Theories at High Energies. 1. General Formalism for Quantum Electrodynamics," *Nucl. Phys.* **B206** (1982) 53.

[63] J. F. Gunion and Z. Kunszt, "Improved Analytic Techniques for Tree Graph Calculations and the $Ggq\bar{q}$ Lepton Anti-Lepton Subprocess," *Phys. Lett.* **B161** (1985) 333.
[64] R. Kleiss and W. J. Stirling, "Spinor Techniques for Calculating $p\bar{p} \to W^{\pm}Z^0 + \text{Jets}$," *Nucl. Phys.* **B262** (1985) 235–262.
[65] N. Arkani-Hamed, F. Cachazo, and J. Kaplan, "What is the Simplest Quantum Field Theory?," *JHEP* **1009** (2010) 016, arXiv:0808.1446 [hep-th].
[66] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*. Wiley Classics Library. New York, John Wiley & Sons Inc., 1978.
[67] R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix*. Cambridge University Press, 1966.
[68] P. Benincasa and F. Cachazo, "Consistency Conditions on the *S*-Matrix of Massless Particles," arXiv:0705.4305 [hep-th].
[69] G. 't Hooft, "A Planar Diagram Theory for Strong Interactions," *Nucl. Phys.* **B72** (1974) 461.
[70] V. P. Nair, "A Current Algebra for Some Gauge Theory Amplitudes," *Phys. Lett.* **B214** (1988) 215.
[71] E. Witten, "Perturbative Gauge Theory as a String Theory in Twistor Space," *Commun. Math. Phys.* **252** (2004) 189–258, arXiv:hep-th/0312171.
[72] M. Bianchi, H. Elvang, and D. Z. Freedman, "Generating Tree Amplitudes in $\mathcal{N}=4$ SYM and $\mathcal{N}=8$ SG," *JHEP* **0809** (2008) 063, arXiv:0805.0757 [hep-th].
[73] J. Drummond, J. Henn, G. Korchemsky, and E. Sokatchev, "Generalized Unitarity for $\mathcal{N}=4$ Super-Amplitudes," *Nucl. Phys.* **B869** (2013) 452–492, arXiv:0808.0491 [hep-th].
[74] R. Britto, F. Cachazo, and B. Feng, "New Recursion Relations for Tree Amplitudes of Gluons," *Nucl. Phys.* **B715** (2005) 499–522, arXiv:hep-th/0412308.
[75] R. Britto, F. Cachazo, B. Feng, and E. Witten, "Direct Proof of Tree-Level Recursion Relation in Yang–Mills Theory," *Phys. Rev. Lett.* **94** (2005) 181602, arXiv:hep-th/0501052.
[76] Z. Bern, L. J. Dixon, and D. A. Kosower, "On-Shell Recurrence Relations for One-Loop QCD Amplitudes," *Phys. Rev.* **D71** (2005) 105013, arXiv:hep-th/0501240.
[77] D. C. Dunbar, J. H. Ettle, and W. B. Perkins, "Augmented Recursion For One-loop Amplitudes," *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 74–79, arXiv:1011.0559 [hep-th].
[78] R. H. Boels, "On BCFW Shifts of Integrands and Integrals," *JHEP* **1011** (2010) 113, arXiv:1008.3101 [hep-th].
[79] S. D. Alston, D. C. Dunbar, and W. B. Perkins, "Complex Factorisation and Recursion for One-Loop Amplitudes," *Phys. Rev.* **D86** (2012) 085022, arXiv:1208.0190 [hep-th].
[80] R. Britto, F. Cachazo, and B. Feng, "Computing One-Loop Amplitudes from the Holomorphic Anomaly of Unitarity Cuts," *Phys. Rev.* **D71** (2005) 025012, arXiv:hep-th/0410179.
[81] S. Caron-Huot, "Loops and Trees," *JHEP* **1105** (2011) 080, arXiv:1007.3224 [hep-ph].
[82] A. P. Hodges, "Twistor Diagrams for All Tree Amplitudes in Gauge Theory: A Helicity-Independent Formalism," arXiv:hep-th/0512336.
[83] M. B. Green, J. H. Schwarz, and L. Brink, "$\mathcal{N}=4$ Yang–Mills and $\mathcal{N}=8$ Supergravity as Limits of String Theories," *Nucl. Phys.* **B198** (1982) 474–492.
[84] C.-N. Yang, "Some Exact Results for the Many Body Problems in One Dimension with Repulsive Delta Function Interaction," *Phys. Rev. Lett.* **19** (1967) 1312–1314.

[85] R. J. Baxter, "Partition Function of the Eight Vertex Lattice Model," *Annals Phys.* **70** (1972) 193–228.
[86] J. Kaplan, "Unraveling $\mathcal{L}_{n,k}$: Grassmannian Kinematics," *JHEP* **1003** (2010) 025, arXiv:0912.0957 [hep-th].
[87] E. I. Buchbinder and F. Cachazo, "Two-Loop Amplitudes of Gluons and Octa-Cuts in $\mathcal{N}=4$ Super Yang–Mills," *JHEP* **0511** (2005) 036, arXiv:hep-th/0506126.
[88] F. Cachazo and D. Skinner, "On the Structure of Scattering Amplitudes in $\mathcal{N}=4$ Super Yang–Mills and $\mathcal{N}=8$ Supergravity," arXiv:0801.4574 [hep-th].
[89] F. Cachazo, "Sharpening The Leading Singularity," arXiv:0803.1988 [hep-th].
[90] F. Cachazo, M. Spradlin, and A. Volovich, "Leading Singularities of the Two-Loop Six-Particle MHV Amplitude," *Phys. Rev.* **D78** (2008) 105022, arXiv:0805.4832 [hep-th].
[91] D. A. Kosower and K. J. Larsen, "Maximal Unitarity at Two Loops," *Phys. Rev.* **D85** (2012) 045017, arXiv:1108.1180 [hep-th].
[92] S. Caron-Huot and K. J. Larsen, "Uniqueness of Two-Loop Master Contours," *JHEP* **1210** (2012) 026, arXiv:1205.0801 [hep-ph].
[93] P. Mastrolia, E. Mirabella, and T. Peraro, "Integrand Reduction of One-Loop Scattering Amplitudes Through Laurent Series Expansion," *JHEP* **1206** (2012) 095, arXiv:1203.0291 [hep-ph].
[94] R. Britto, F. Cachazo, and B. Feng, "Generalized Unitarity and One-Loop Amplitudes in $\mathcal{N}=4$ Super-Yang–Mills," *Nucl. Phys.* **B725** (2005) 275–305, arXiv:hep-th/0412103.
[95] W. L. van Neerven and J. A. M. Vermaseren, "Large Loop Integrals," *Phys. Lett.* **B137** (1984) 241.
[96] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, "One-Loop *n*-Point Gauge Theory Amplitudes, Unitarity and Collinear Limits," *Nucl. Phys.* **B425** (1994) 217–260, arXiv:hep-ph/9403226.
[97] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, "Fusing Gauge Theory Tree Amplitudes into Loop Amplitudes," *Nucl. Phys.* **B435** (1995) 59–101, arXiv:hep-ph/9409265.
[98] Z. Bern, J. Rozowsky, and B. Yan, "Two-Loop Four-Gluon Amplitudes in $\mathcal{N}=4$ Super Yang–Mills," *Phys. Lett.* **B401** (1997) 273–282, arXiv:hep-ph/9702424.
[99] Z. Bern, L. J. Dixon, and V. A. Smirnov, "Iteration of Planar Amplitudes in Maximally Supersymmetric Yang–Mills Theory at Three Loops and Beyond," *Phys. Rev.* **D72** (2005) 085001, arXiv:hep-th/0505205.
[100] Z. Bern, J. Carrasco, H. Johansson, and D. Kosower, "Maximally Supersymmetric Planar Yang–Mills Amplitudes at Five Loops," *Phys. Rev.* **D76** (2007) 125020, arXiv:0705.1864 [hep-th].
[101] G. Ossola, C. G. Papadopoulos, and R. Pittau, "CutTools: A Program Implementing the OPP Reduction Method to Compute One-Loop Amplitudes," *JHEP* **0803** (2008) 042, arXiv:0711.3596 [hep-ph].
[102] I. M. Gelfand, R. M. Goresky, R. D. MacPherson, and V. V. Serganova, "Combinatorial Geometries, Convex Polyhedra, and Schubert Cells," *Adv. in Math.* **63** (1987) no. 3, 301–316.
[103] N. E. Mnëv, "The Universality Theorems on the Classification Problem of Configuration Varieties and Convex Polytope Varieties," in *Topology and Geometry—Rohlin Seminar*, vol. 1346 of *Lecture Notes in Math.*, pp. 527–543. Springer, Berlin, 1988.
[104] K. Rietsch, "An Algebraic Cell Decomposition of the Nonnegative Part of a Flag Variety," *J. Algebra* **213** (1999) no. 1, 144–154, arXiv:alg-geom/9709035.
[105] L. K. Williams, "Shelling Totally Nonnegative Flag Varieties," *J. Reine Angew. Math.* **609** (2007) 1–21, arXiv:math/0509129 [math.CO].

[106] L. K. Williams, "Enumeration of Totally Positive Grassmann Cells," *Adv. Math.* **190** (2005) no. 2, 319–342, arXiv:math/0307271 [math.CO].
[107] S. J. Parke and T. R. Taylor, "An Amplitude for n-Gluon Scattering," *Phys. Rev. Lett.* **56** (1986) 2459.
[108] R. Penrose, "Twistor Algebra," *J. Math. Phys.* **8** (1967) 345.
[109] R. Penrose and M. A. H. MacCallum, "Twistor Theory: An Approach to the Quantization of Fields and Space-Time," *Phys. Rept.* **6** (1972) 241–316.
[110] R. Penrose, "Twistor Quantization and Curved Space-Time," *Int. J. Theor. Phys.* **1** (1968) 61–99.
[111] E. H. Kronheimer and R. Penrose, "On the Structure of Causal Spaces," *Proc. Cambridge Phil. Soc.* **63** (1967) 481–501.
[112] R. Penrose, "The Central Programme of Twistor Theory," *Chaos Solitons Fractals* **10** (1999) 581–611.
[113] A. Ferber, "Supertwistors and Conformal Supersymmetry," *Nucl. Phys.* **B132** (1978) 55.
[114] J. M. Drummond, J. M. Henn, and J. Plefka, "Yangian Symmetry of Scattering Amplitudes in $\mathcal{N} = 4$ Super Yang–Mills Theory," *JHEP* **05** (2009) 046, arXiv:0902.2987 [hep-th].
[115] J. Drummond and L. Ferro, "The Yangian Origin of the Grassmannian Integral," *JHEP* **1012** (2010) 010, arXiv:1002.4622 [hep-th].
[116] J. Drummond and L. Ferro, "Yangians, Grassmannians and T-duality," *JHEP* **1007** (2010) 027, arXiv:1001.3348 [hep-th].
[117] N. Beisert, J. Henn, T. McLoughlin, and J. Plefka, "One-Loop Superconformal and Yangian Symmetries of Scattering Amplitudes in $\mathcal{N} = 4$ Super Yang–Mills," *JHEP* **04** (2010) 085, arXiv:1002.1733 [hep-th].
[118] J. Drummond, J. Henn, V. Smirnov, and E. Sokatchev, "Magic Identities for Conformal Four-Point Integrals," *JHEP* **0701** (2007) 064, arXiv:hep-th/0607160.
[119] L. F. Alday and J. M. Maldacena, "Gluon Scattering Amplitudes at Strong Coupling," *JHEP* **06** (2007) 064, arXiv:0705.0303 [hep-th].
[120] L. F. Alday and J. Maldacena, "Comments on Gluon Scattering Amplitudes via AdS/CFT," *JHEP* **11** (2007) 068, arXiv:0710.1060 [hep-th].
[121] A. Brandhuber, P. Heslop, and G. Travaglini, "MHV Amplitudes in $\mathcal{N} = 4$ Super Yang–Mills and Wilson Loops," *Nucl. Phys.* **B794** (2008) 231–243, arXiv:0707.1153 [hep-th].
[122] J. M. Drummond, G. P. Korchemsky, and E. Sokatchev, "Conformal Properties of Four-Gluon Planar Amplitudes and Wilson loops," *Nucl. Phys.* **B795** (2008) 385–408, arXiv:0707.0243 [hep-th].
[123] J. M. Drummond, J. Henn, G. P. Korchemsky, and E. Sokatchev, "On Planar Gluon Amplitudes/Wilson Loops Duality," *Nucl. Phys.* **B795** (2008) 52–68, arXiv:0709.2368 [hep-th].
[124] J. M. Drummond, J. Henn, G. P. Korchemsky, and E. Sokatchev, "The Hexagon Wilson Loop and the BDS Ansatz for the Six-Gluon Amplitude," *Phys. Lett.* **B662** (2008) 456–460, arXiv:0712.4138 [hep-th].
[125] J. M. Drummond, J. Henn, G. P. Korchemsky, and E. Sokatchev, "Hexagon Wilson Loop = Six-Gluon MHV Amplitude," *Nucl. Phys.* **B815** (2009) 142–173, arXiv:0803.1466 [hep-th].
[126] Z. Bern *et al.*, "The Two-Loop Six-Gluon MHV Amplitude in Maximally Supersymmetric Yang–Mills Theory," *Phys. Rev.* **D78** (2008) 045007, arXiv:0803.1465 [hep-th].
[127] L. F. Alday and J. Maldacena, "Null Polygonal Wilson Loops and Minimal Surfaces in Anti-de-Sitter Space," *JHEP* **11** (2009) 082, arXiv:0904.0663 [hep-th].

[128] A. Hodges, "Eliminating Spurious Poles from Gauge-Theoretic Amplitudes," *JHEP* **1305** (2013) 135, arXiv:0905.1473 [hep-th].
[129] G. Korchemsky and E. Sokatchev, "Superconformal Invariants for Scattering Amplitudes in $\mathcal{N}=4$ SYM Theory," *Nucl. Phys.* **B839** (2010) 377–419, arXiv:1002.4625 [hep-th].
[130] J. Drummond, J. Henn, G. Korchemsky, and E. Sokatchev, "Dual Superconformal Symmetry of Scattering Amplitudes in $\mathcal{N}=4$ Super Yang–Mills Theory," *Nucl. Phys.* **B828** (2010) 317–374, arXiv:0807.1095 [hep-th].
[131] N. Arkani-Hamed, J. Bourjaily, F. Cachazo, and J. Trnka, "Unification of Residues and Grassmannian Dualities," *JHEP* **1101** (2011) 049, arXiv:0912.4912 [hep-th].
[132] N. Arkani-Hamed, J. Bourjaily, F. Cachazo, and J. Trnka, "Local Spacetime Physics from the Grassmannian," *JHEP* **1101** (2011) 108, arXiv:0912.3249 [hep-th].
[133] M. Bullimore, L. Mason, and D. Skinner, "Twistor-Strings, Grassmannians and Leading Singularities," *JHEP* **1003** (2010) 070, arXiv:0912.0539 [hep-th].
[134] S. K. Ashok and E. Dell'Aquila, "On the Classification of Residues of the Grassmannian," *JHEP* **1110** (2011) 097, arXiv:1012.5094 [hep-th].
[135] M. Staudacher, "Review of AdS/CFT Integrability, Chapter III.1: Bethe Ansätze and the R-Matrix Formalism," *Lett. Math. Phys.* **99** (2012) 191–208, arXiv:1012.3990 [hep-th].
[136] M. S. Bianchi, M. Leoni, and S. Penati, "An All Order Identity between ABJM and $\mathcal{N}=4$ SYM Four-Point Amplitudes," *JHEP* **1204** (2012) 045, arXiv:1112.3649 [hep-th].
[137] M. S. Bianchi, M. Leoni, A. Mauri, S. Penati, and A. Santambrogio, "Scattering Amplitudes/Wilson Loop Duality In ABJM Theory," *JHEP* **1201** (2012) 056, arXiv:1107.3139 [hep-th].
[138] M. S. Bianchi, M. Leoni, A. Mauri, S. Penati, and A. Santambrogio, "Scattering in ABJ theories," *JHEP* **1112** (2011) 073, arXiv:1110.0738 [hep-th].
[139] M. S. Bianchi, M. Leoni, A. Mauri, S. Penati, and A. Santambrogio, "One Loop Amplitudes In ABJM," *JHEP* **1207** (2012) 029, arXiv:1204.4407 [hep-th].
[140] T. Bargheer, F. Loebbert, and C. Meneghelli, "Symmetries of Tree-Level Scattering Amplitudes in $\mathcal{N}=6$ Superconformal Chern–Simons Theory," *Phys. Rev.* **D82** (2010) 045016, arXiv:1003.6120 [hep-th].
[141] S. Lee, "Yangian Invariant Scattering Amplitudes in Supersymmetric Chern–Simons Theory," *Phys. Rev. Lett.* **105** (2010) 151603, arXiv:1007.4772 [hep-th].
[142] Y.-t. Huang and A. E. Lipstein, "Dual Superconformal Symmetry of $\mathcal{N}=6$ Chern–Simons Theory," *JHEP* **1011** (2010) 076, arXiv:1008.0041 [hep-th].
[143] D. Gang, Y.-t. Huang, E. Koh, S. Lee, and A. E. Lipstein, "Tree-level Recursion Relation and Dual Superconformal Symmetry of the ABJM Theory," *JHEP* **1103** (2011) 116, arXiv:1012.5032 [hep-th].
[144] A. Agarwal, N. Beisert, and T. McLoughlin, "Scattering in Mass-Deformed $\mathcal{N}\geq 4$ Chern–Simons Models," *JHEP* **0906** (2009) 045, arXiv:0812.3367 [hep-th].
[145] T. Bargheer, N. Beisert, F. Loebbert, T. McLoughlin, N. Beisert, *et al.*, "Conformal Anomaly for Amplitudes in $\mathcal{N}=6$ Superconformal Chern–Simons Theory," *J. Phys.* **A45** (2012) 475402, arXiv:1204.4406 [hep-th].
[146] A. Brandhuber, G. Travaglini, and C. Wen, "A Note on Amplitudes in $\mathcal{N}=6$ Superconformal Chern–Simons Theory," *JHEP* **1207** (2012) 160, arXiv:1205.6705 [hep-th].
[147] A. Brandhuber, G. Travaglini, and C. Wen, "All One-Loop Amplitudes in $\mathcal{N}=6$ Superconformal Chern–Simons Theory," *JHEP* **1210** (2012) 145, arXiv:1207.6908 [hep-th].

[148] M. Kontsevich, "Deformation Quantization of Poisson Manifolds," *Lett. Math. Phys.* **66** (2003) no. 3, 157–216, arXiv:q-alg/9709040.
[149] S. Caron-Huot, "Notes on the Scattering Amplitude / Wilson Loop Duality," *JHEP* **1107** (2011) 058, arXiv:1010.1167 [hep-th].
[150] L. Mason and D. Skinner, "The Complete Planar *S*-Matrix of $\mathcal{N}=4$ SYM as a Wilson Loop in Twistor Space," *JHEP* **12** (2010) 018, arXiv:1009.2225 [hep-th].
[151] M. Bullimore and D. Skinner, "Holomorphic Linking, Loop Equations and Scattering Amplitudes in Twistor Space," arXiv:1101.1329 [hep-th].
[152] L. F. Alday, B. Eden, G. P. Korchemsky, J. Maldacena, and E. Sokatchev, "From Correlation Functions to Wilson Loops," *JHEP* **1109** (2011) 123, arXiv:1007.3243 [hep-th].
[153] B. Eden, G. P. Korchemsky, and E. Sokatchev, "From Correlation Functions to Scattering Amplitudes," *JHEP* **1112** (2011) 002, arXiv:1007.3246 [hep-th].
[154] B. Eden, G. P. Korchemsky, and E. Sokatchev, "More on the Duality Correlators/Amplitudes," *Phys. Lett.* **B709** (2012) 247–253, arXiv:1009.2488 [hep-th].
[155] B. Eden, P. Heslop, G. P. Korchemsky, and E. Sokatchev, "The Super-Correlator/Super-Amplitude Duality: Part I," *Nucl. Phys.* **B869** (2013) 329–377, arXiv:1103.3714 [hep-th].
[156] B. Eden, P. Heslop, G. P. Korchemsky, and E. Sokatchev, "The Super-Correlator/Super-Amplitude Duality: Part II," *Nucl. Phys.* **B869** (2013) 378–416, arXiv:1103.4353 [hep-th].
[157] P. C. Schuster and N. Toro, "Constructing the Tree-Level Yang–Mills *S*-Matrix Using Complex Factorization," *JHEP* **0906** (2009) 079, arXiv:0811.3207 [hep-th].
[158] S. He and H.-b. Zhang, "Consistency Conditions on *S*-Matrix of Spin 1 Massless Particles," *JHEP* **1007** (2010) 015, arXiv:0811.3210 [hep-th].
[159] J. M. Drummond and J. M. Henn, "All Tree-Level Amplitudes in $\mathcal{N}=4$ SYM," *JHEP* **04** (2009) 018, arXiv:0808.2475 [hep-th].
[160] J. L. Bourjaily, "Efficient Tree-Amplitudes in $\mathcal{N}=4$: Automatic BCFW Recursion in MATHEMATICA," arXiv:1011.2447 [hep-ph].
[161] N. Arkani-Hamed, F. Cachazo, C. Cheung, and J. Kaplan, "The *S*-Matrix in Twistor Space," *JHEP* **1003** (2010) 110, arXiv:0903.2110 [hep-th].
[162] Z. Bern, "String Based Perturbative Methods for Gauge Theories," arXiv:hep-ph/9304249.
[163] Z. Bern and A. Morgan, "Supersymmetry Relations Between Contributions to One-Loop Gauge Boson Amplitudes," *Phys. Rev.* **D49** (1994) 6155–6163, arXiv:hep-ph/9312218.
[164] A. Kotikov and L. Lipatov, "DGLAP and BFKL Equations in the $\mathcal{N}=4$ Supersymmetric Gauge Theory," *Nucl. Phys.* **B661** (2003) 19–61, arXiv:hep-ph/0208220.
[165] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, and J. Trnka, "Local Integrals for Planar Scattering Amplitudes," *JHEP* **1206** (2012) 125, arXiv:1012.6032 [hep-th].
[166] A. Brandhuber, O. Gurdogan, R. Mooney, G. Travaglini, and G. Yang, "Harmony of Super Form Factors," *JHEP* **1110** (2011) 046, arXiv:1107.5067 [hep-th].
[167] L. F. Alday, D. Gaiotto, and Y. Tachikawa, "Liouville Correlation Functions from Four-Dimensional Gauge Theories," *Lett. Math. Phys.* **91** (2010) 167–197, arXiv:0906.3219 [hep-th].
[168] L. F. Alday, "Review of AdS/CFT Integrability, Chapter V.3: Scattering Amplitudes at Strong Coupling," *Lett. Math. Phys.* **99** (2012) 507–528, arXiv:1012.4003 [hep-th].
[169] N. Arkani-Hamed, A. G. Cohen, D. B. Kaplan, A. Karch, and L. Motl, "Deconstructing $(2,0)$ and Little String Theories," *JHEP* **0301** (2003) 083, arXiv:hep-th/0110146.

[170] C. Cheung and D. O'Connell, "Amplitudes and Spinor-Helicity in Six Dimensions," *JHEP* **0907** (2009) 075, arXiv:0902.0981 [hep-th].
[171] We thank Yu-tin Huang for discussions on this point.
[172] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. Hodges, and J. Trnka, "A Note on Polytopes for Scattering Amplitudes," *JHEP* **1204** (2012) 081, arXiv:1012.6030 [hep-th].

Index

3-particle helicity amplitudes, 10–13
3-particle superamplitude $\mathcal{A}_3^{(1)}$, 41
  Grassmannian representation of, 42
3-particle superamplitude $\mathcal{A}_3^{(2)}$, 42
  Grassmannian representation of, 42
5-bracket, 109, 121
*d*LIPS, 6

amalgamation
  of on-shell diagrams, 45–49
  of positroid configurations, 72

bad double-crossing, 35
BCFW all-loop recursion relations, 22, **155**
  diagrammatic proof of, 156–159
BCFW one-loop recursion
  number of terms, 171
BCFW tree-level recursion, 160
  number of terms, 164
BCFW-bridge
  adding a bridge to an on-shell diagram, 18, **162**
    action on permutation labels, 33
BCFW-bridge construction, 33–36
  generation of canonical coordinates, 66
  lexicographic scheme, **34**
BCFW-bridge coordinates, 66, 67
  accessibility of boundary configurations, 82
  detailed example of, 68, 75
  positivity of, 73–76
boundary measurements, 49–52
  in terms of edge variables, 50
  in terms of face variables, 52
bubble-reduction, 27
  transformation of edge variables, 55
  transformation of face variables, 55

canonical coordinates, 49–52, 73
  approaching boundaries in, 81–82
  edge variables, 49
    for 3-particle vertices, 49
    transformation under bubble-reduction, 55
  face variables, 45, 48, 49
    for 3-particle vertices, 48
    transformation under bubble-reduction, 55
    transformation under merge/unmerge operation, 53
    transformation under square move, 54
  positive canonical coordinates, 73
  transformations of, 52
canonical volume form on $G_+(k,n)$, 83
  in terms of edge variables, 51
  in terms of face variables, 49
conformal symmetry
  momentum-space generators of, 90
  twistor-space generators of, 91
Cramer's rule, 38
  application to 2-vectors, 41
  application to 4-vectors, 96

decorated permutations, 31, 63
distinguished boundary elements $[\hat{\partial}^k](C)$, 102
*d*log-measure, 69

edge variables, 49
  and boundary measurements, 50
  for 3-particle vertices, 49
  positroid volume form in terms of, 51
  transformation under bubble-reduction, 55
Eulerian poset, 80

face variables, 45, 48
  and boundary measurements, 52
  for 3-particle vertices, 48
  positroid volume form in terms of, 49
  transformation under bubble-reduction, 55
  transformation under merge/unmerge operation, 53
  transformation under square move, 54
factorization channel, 8
factorization channels, 21, 23, 155, 156
fermionic
  δ-functions, 16
  parameters, 15

five-bracket
  identity involving, 110
forward-limits, 155, 156
  in theories with $\mathcal{N} < 4$, 178
four-mass box, 105, 110, 122
  identity involving, 124
  kinematical support of, 105

gauge-fixing
  of a representative element of $G(k,n)$, 37
$G(k,n)$, 37
$G_+(k,n)$, *see* positive Grassmannian, $G_+(k,n)$
Grassmann coherent state, 14
Grassmannian necklace, 65
Grassmannian of k-planes in n dimensions, *see* $G(k,n)$

hooks (in the $r[a;b]$ table), 63

invariant top-form, $\mathbf{\Omega}^{\text{top}}$, on $G_+(k,n)$, 83
  in terms of canonical coordinates, 83
  in terms of matrix coordinates, 83
inverse-soft constructibility, 162
  inverse-soft constructible recursion scheme, 163
  non-inverse-soft constructible example, 162
inverse-soft factors, 161

juggling patterns, 64
  hooks, 63

kinematical support of on-shell diagrams, 101–106

leading singularities, 9, 84
  composite leading singularities, 56
left-right path permutation labels, 31
little group
  definition of, 6
$\mathcal{L}_{n,k}$, 108
Lorentz transformations, 40, 90, 131, 180
Lorentz-invariant phase space, 6

matrix coordinates for $C \in G(k,n)$, 83
merge/unmerge operation, 24–27, 32, 52, 141, 172
  relation to Yang–Baxter, 129
  transformation of face variables, 52, 53
merge/unmerge operation, 139
MHV: *maximal helicity violating*, 15
momentum conservation, 7, 40–42, 90, 107
  for 3-particle kinematics, 11
momentum-twistor variables, 92–96

observable states, 5
on-shell diagrams
  amalgamation of, 45
    direct product, 46
    projections, 46
  bad double-crossing, 35
  boundaries of, 78–79
  boundary measurements, 49–52
  construction of, 5–20
  dimension of, 45
  Grassmannian representations of, 41–45
  introduction to, 8–18
  left-right path permutations, 31
  physical equivalences of, 23–28
  physical singularities of, 77–79
  reducibility of, 27, 32, 36
on-shell differential forms
  Grassmannian representations of, 41–45
  in terms of canonical coordinates, 88
  in terms of matrix coordinates, 88
  introduction to, 8–18
  kinematical support of, 101–106
  twistor-space representations of, 92
on-shell functions
  construction of, 5
on-shell particles
  definition of, 5
orthogonal complement $C^{\perp}$ of $C \in G(k,n)$, 38
  components of, 39

Parke–Taylor formula, 89
permutation labels
  equivalence between positroid geometry and on-shell diagrams labels, 66
  for on-shell diagrams, 31
  for positroid configurations, 64
Plücker relations, 38
positive (canonical) coordinates, 73
  for generic positroid configurations, 73
    relation to BCFW-bridge coordinates, 74
positive configurations, *see* positroid configurations
positive Grassmannian, $G_+(k,n)$, 60–76
  configurations in $G_+(1,n)$, 69
  configurations in $G_+(3,6)$, 61
  configurations in $G_+(4,8)$, 62
positive subspace of $\mathbb{RP}^n$, 70
positroid boundaries
  distinguished boundary elements $[\partial^k](C)$, 102
positroid configurations, 61, 67, 69, 73, 74
  boundaries of, 78
  dimensionality of, 64
positroid stratification, 61, 65, 69, 72
  canonical volume form, 83
  complete stratification of $G_+(2,4)$, 81
  covering relations, 78
  graphical representation of $G_+(2,4)$, 81

$r[a;b] \equiv \text{rank}\{c_a, \ldots, c_b\}$, 62

scattering amplitudes
  3-particle helicity amplitudes, 10
  BCFW all-loop recursion relations, 22, **155**
  differential equation for, 155
  MHV amplitudes
    4-particle, one-loop, 26, 165–168
    n-particle, one-loop, 170–171
    4-particle, tree-level, 20, 160
    5-particle, tree-level, 160
    6-particle, tree-level, 161

scattering amplitudes (*cont.*)
n-particle, tree-level, 89
$\overline{\text{MHV}}$ amplitudes
5-particle, tree-level, 160
6-particle, tree-level, 161
NMHV amplitudes
6-particle, tree-level, 107, 108, 163, 164
singularities of, 9
4-particle tree amplitude, 156
transcendentality of, 173–177
S-matrix, 29, 30
in (1+1) dimensions, 29, 128
spinor-helicity variables, 6
Grassmannian geometry of, 40–41
spurious poles, 158, 159
square move, 24–27, 32, 33, 52, 139, 141, 143, 172
action on face variables, 53–55
and cluster coordinates, 147, 148
and Seiberg duality, 145
in ABJM theories, 134
relation to Yang–Baxter, 129
transformation of face variables, 54

transcendentality of loop-amplitudes, 173–177
twisted cyclic symmetry, 70
twistor variables
definition of, 90

wick contractions, 13

Yang–Baxter relation, 30
on-shell avatar of, 127–130