arXiv:1708.03872v1 [hep-ph] 13 Aug 2017

TASI Lectures on Scattering Amplitudes

Clifford Cheung

Walter Burke Institute for Theoretical Physics
California Institute of Technology, Pasadena, CA 91125

These lectures are a brief introduction to scattering amplitudes. We begin with a review of basic kinematical concepts like the spinor helicity formalism, followed by a tutorial on bootstrapping tree-level scattering amplitudes. Afterwards, we discuss on-shell recursion relations and soft theorems, emphasizing their broad applicability to gravity, gauge theory, and effective field theories. Lastly, we report on some of the new field theoretic structures which have emerged from the on-shell picture, focusing primarily on color-kinematics duality.

Contents

1   Introductory Remarks

1.1   Symmetry and Redundancy

Quantum field theory (QFT) is a tremendously powerful tool. It boasts applications across a range of fields, including particle physics, string theory, cosmology, statistical mechanics, condensed matter theory, and even economics. In short, QFT delivers a *quantitive* description of the totality of the cosmos down to the shortest distance scales probed by humankind, all while enriching you financially.

But what actually defines a QFT? According to most textbook treatments, the answer is *the action*. The reasons for this are practical. First of all, the action formulation of QFT is algorithmic, automating much of the heavy lifting needed to translate high-minded theoretical structures into physical predictions. Second, the action is borne of our classical description of the world, so it preserves—for better or worse—those natural intuitions.

Nevertheless, the luxuries of the action come at a cost, which is *redundancy*. In the context of electromagnetism, this appears as the invariance of the action under gauge transformations of the photon,

$$A_\mu \to A_\mu + \partial_\mu \theta. \tag{1.1}$$

The action describing interacting gravitons is similarly invariant under diffeomorphisms, which at linear order take the form,

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \theta_\nu + \partial_\nu \theta_\mu. \tag{1.2}$$

Here the transformation parameters θ and θ_μ are *arbitrary* functions of spacetime with compact support.

At face value, the appearance of an entire function's worth of freedom attests to an infinite amount of "symmetry". For an apples to apples comparison, contrast this with a global Lorentz transformation, which is parameterized by six numbers for three rotations and three boosts. This would naively suggest that gauge and diffeomorphism "symmetries" are infinitely more potent than their anemic global siblings. However, as foreshadowed by the above quotation marks, this is exactly the opposite of the truth. Gauge and diffeomorphism invariance are actually redundancies of description created by physicists to conveniently describe certain QFTs.

To understand the purpose of this contrivance, recall the Wigner classification [1, 2], which is a sort of physicist's inventory of "what kind of

stuff" can exist in the universe even in principle. That such a fundamental question—once thought to be squarely within the jurisdiction of philosophical navel-gazing—is answerable with absolute mathematical precision is one of the great achievements of modern physics. In any case, the Wigner construction simply amounts to enumerating all possible irreducible unitary representations of the Poincare group.

In four dimensions, these arguments imply that a massless particle with intrinsic spin has exactly two polarizations. Unfortunately, our preference for manifest Lorentz invariance compels us to describe the photon and graviton with quantum fields, $A_\mu(x)$ and $h_{\mu\nu}(x)$, whose many indices beget four and ten degrees of freedom, respectively. As these redundant modes are a necessary evil of manifest Lorentz covariance, we have little choice but to introduce gauge and diffeomorphism *redundancies* to stamp out the many degrees of freedom overzealously introduced.

The above examples suggest that this redundancy is an affliction of theories with spin. However, it is actually *endemic* to the action principle. For instance, consider a *scalar* field theory with the general Lagrangian[1],

$$\mathcal{L} = \frac{1}{2}K(\phi)\partial_\mu\phi\partial^\mu\phi, \tag{1.3}$$

where the field dependent kinetic term is

$$K(\phi) = 1 + \lambda_1\phi + \frac{1}{2!}\lambda_2\phi^2 + \frac{1}{3!}\lambda_3\phi^3 + \frac{1}{4!}\lambda_4\phi^4 + \ldots \tag{1.4}$$

Given the complicated and arbitrary tower of interactions defined in Eq. (1.3), one naturally anticipates a similarly convoluted S-matrix.

To test this intuition, consider the simplest physical observable: the four-particle tree-level scattering amplitude. A simple calculation yields

$$A_4 \propto \sum_{i \neq j} p_i p_j \propto s + t + u = 0, \tag{1.5}$$

where the usual Mandelstam invariants [3] are

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2 \tag{1.6}$$
$$t = (p_1 + p_4)^2 = (p_2 + p_3)^2 \tag{1.7}$$
$$u = (p_1 + p_3)^2 = (p_2 + p_4)^2. \tag{1.8}$$

---

[1] We employ mostly minus conventions where $\eta_{\mu\nu} = \mathrm{diag}(+, -, \ldots, -)$.

Here it was absolutely crucially to impose physical external kinematics, *i.e.* momenta which conserve total momentum,

$$\sum_i p_i = 0, \tag{1.9}$$

and satisfy the on-shell conditions,

$$p_i^2 = 0. \tag{1.10}$$

So the four-particle amplitude actually vanishes. This effect persists—for example, the fourteen-particle amplitude also vanishes, albeit through the diabolical cancellation of upwards of five trillion Feynman diagrams[2].

The origin of this conspiracy is that the action in Eq. (1.3) is secretly that of a *free scalar*. In particular, the S-matrix is invariant under arbitrary field redefinitions of the form [4, 5],

$$\phi \to f(\phi), \tag{1.11}$$

subject to the condition $f'(0) = 1$ so that the weak field limit interpolates into a canonically normalized field. For a suitably chosen function $f(\phi)$, one can then transform the free scalar action into Eq. (1.3). So the S-matrix is invariant under a *non-symmetry* of the action! In hindsight, this freedom of field basis should not be all that surprising. After all, the path integral formulation of QFT treats the field as an *integration variable* and any integral is invariant under a proper change of variables.

We thus arrive at the bottom line: the textbook formulation of QFT is plagued by a redundancy that undercuts any fundamental meaning to contents of the action. By applying any old field redefinition, we can map one action into an infinite set of *different* actions describing *identical* physics.

This infinity-to-one mapping has its disadvantages. First of all, these redundancies spawn a correspondingly redundant set of Feynman diagram representations of the same S-matrix. For theories with gauge and diffeomorphism invariance, the equivalence of these representations is manifested by the Ward identities. To satisfy the Ward identities the resulting Feynman diagram expressions must be immensely complicated, often exceeding the capabilities of a non-computer-assisted being even for simple physical processes. Second and more importantly, a poor choice of field basis may obscure or *altogether conceal* certain underlying structures of the theory.

---

[2] See Appendix A for the counting of Feynman diagrams in a general QFT.

1.2   What Defines a Theory?

So what defines a theory, if not the action? To answer this question, consider the following Gedanken sociology experiment. Your friend resides in another pocket of the multiverse whose entire field content is a set of massless fields which are mutually interacting. After a prolonged and ad hominem exchange about who is or is not a Boltzmann brain, they challenge you in exasperation: if you are so smart, can you compute their S-matrix from first principles?

For scalar fields, your first emotion might be one of relief—after all scalar Feynman diagrams are easy to compute. However, without any additional information, the interactions are various and sundry,

$$\mathcal{L}_{\text{scalar}} = \frac{1}{2}\sum_a \partial_\mu \phi_a \partial^\mu \phi_a + \frac{1}{3!}\sum_{abc} \omega_{abc}\phi_a\phi_b\phi_c + \frac{1}{4!}\sum_{abcd} \lambda_{abcd}\phi_a\phi_b\phi_c\phi_d. \quad (1.12)$$

So we need to specify all of the coupling constants ω_{abc} and λ_{abcd} to even define the action. In other words, the theory might be simple with respect to Feynman diagrams, but it is not predictive in the slightest.

If, on the other hand, the fields are vectors, then there is almost no freedom: the *only* theory permitted is Yang-Mills (YM) theory,

$$\mathcal{L}_{\text{vector}} = -\frac{1}{4g^2}\sum_a F_{a\mu\nu}F_a^{\mu\nu}. \quad (1.13)$$

As we will soon see in explicit detail, this is not simply an educated guess. One can *prove* that there are no interacting renormalizable theories of massless vectors aside from YM. In this approach, the fact that the structure constants f_{abc} of the gauge group are antisymmetric and satisfy the Jacobi identities is an *output*.

Similar logic implies that there are no self-consistent renormalizable interactions of a massless tensor. Instead, the leading allowed interactions are nonrenormalizable and coincide precisely with the Einstein-Hilbert action,

$$\mathcal{L}_{\text{tensor}} = -\frac{1}{16\pi G}\sqrt{-g}\,R. \quad (1.14)$$

So the low-energy dynamics of a massless tensor are controlled by a single parameter which is Newton's constant.

Similar statements apply to fermions. For example, spin $1/2$ fields are typically less constrained in their interactions while a spin $3/2$ field can only interact consistently via supergravity [6].

This little parable serves to illustrate a trivial but important point: the simplicity of a theory hinges on how little must be specified in order to deduce its S-matrix. Oftentimes, the mere presence of higher spin particles can be enough to fix the structure of interactions. In some special circumstances, this can also be true for scalars, *e.g.* the pion. As a general principle, the simplest QFTs are so constrained that *were they any more constrained they would not even exist.*

Among amplitudes practitioners, there is a growing consensus that $\mathcal{N} = 4$ supersymmetric Yang-Mills (SYM) theory is the simplest QFT. To the uninitiated, this claim may appear ludicrous. After all, $\mathcal{N} = 4$ SYM is comprised of numerous fields interacting through a lengthy component action. However, this criticism is akin to bemoaning the byzantine complexity of the quantum harmonic oscillator because Hermite polynomials are hard to manipulate and there are a lot of them. Hopefully, by the end of these lectures you will understand that these naive intuitions are misguided. A theory like $\mathcal{N} = 4$ SYM is simple because its full perturbative S-matrix is fixed by symmetry. The fact that the action seems complicated is merely an indication that there is a better way to compute!

The modern amplitudes program is an intellectual descendant of the fabled "Feynman method" of calculation, which is to *i*) define the problem, *ii*) think very hard, and *iii*) write down the answer while being Richard Feynman. Notably, mileage tends to vary in the critical last step. A more practical approach is to construct a *general ansatz* for the S-matrix and sculpt out the correct answer from simple physical criteria:

- **Dimensional Analysis.**

  Scattering amplitudes should have mass dimension consistent with the dimensionality of the coupling constants in the theory. While this can be trivially gleaned from an action, it is still useful in practice.

- **Lorentz Invariance.**

  Scattering amplitudes should be Lorentz invariant. For example, a four-particle amplitude of scalars is a function of s, t, u. When there are particles with spin, the amplitude should also be covariant under the little group, which is a close cousin of Lorentz invariance.

- **Locality.**

  Scattering amplitudes should have kinematic singularities which are consistent with factorization and unitarity. These singularities encode underlying locality of the theory. For example, a four-particle amplitude can have poles like $1/s$ but not $1/s^2$.

Remarkably, from these basic principles one can bootstrap a tremendous amount of physics.

2 Crash Course in Kinematics

Kinematics is at the heart of the modern S-matrix program. A staple of every introductory course in QFT, this subject is typically given short shrift in the mad dash to dynamics. However, as we will soon see, any sharp distinction between kinematics and dynamics is very much artificial. Many scattering amplitudes are actually *uniquely* fixed by kinematic constraints like Lorentz invariance and momentum conservation.

In the textbook formulation of QFT, kinematic data is characterized more or less exclusively by momentum vectors,

$$p_\mu = (p_0, \vec{p}), \tag{2.1}$$

where throughout these lectures we take the convention that all particles are *incoming*. Over the years, the study of scattering amplitudes program has revealed a menagerie of alternative variables: spinor helicity [7, 8, 9, 10] , twistors [11], and momentum twistors [12], just to name a few. Of course, these all describe the same underlying physics. However, by recasting expressions into the appropriate variables, one can achieve massive simplifications which reveal otherwise invisible structures. In short, never bring a knife to a gun fight because they have different symmetries.

2.1 Spinor Helicity

Spinor helicity variables are a tremendously powerful tool for representing on-shell kinematics. While our discussion here will center exclusively on spinor helicity formalism in four dimensions, generalizations exist for lower [13] and higher [14, 15] dimensions.

In a nutshell, the spinor helicity formalism maps the components of a four-vector into those of a two-by-two matrix via

$$p_{\alpha\dot\alpha} = p_\mu \sigma^\mu_{\alpha\dot\alpha} = \begin{pmatrix} p_0 + p_3 & p_1 - ip_2 \\ p_1 + ip_2 & p_0 - p_3 \end{pmatrix}, \qquad (2.2)$$

where $\sigma^\mu = (1, \vec\sigma)$ is a four-vector of Pauli matrices and the undotted and dotted indices transform under the usual spinor representations of the Lorentz group. The only Lorentz invariant quantity which can be constructed from $p_{\alpha\dot\alpha}$ is its determinant,

$$\det(p) = p^\mu p_\mu = 0, \qquad (2.3)$$

which vanishes for a massless on-shell particle.

The spinor helicity formalism has been generalized to include massive particles [16, 17, 18]. However, the honest truth is that theories with massless particles are unambiguously simpler than their massive counterparts. This fact transcends the context of scattering amplitudes. Even in the action formulation, particles like gluons, gravitons, chiral fermions, or Nambu-Goldstone bosons (NGBs) are all massless as a consequence of symmetries which also stringently constrain their permitted interactions. This is the reason why we so often define massive particles as a *controlled perturbation* away from the massless limit, *e.g.* in the case of soft breaking of supersymmetry or chiral symmetry.

Returning to Eq. (2.3), we see that $p_{\alpha\dot\alpha}$ has vanishing determinant. Since $p_{\alpha\dot\alpha}$ is nonzero, it is a two-by two matrix of at most rank one, which without loss of generality can be written as the outer product of two two-component objects which are called spinors,

$$p_{\alpha\dot\alpha} = \lambda_\alpha \tilde\lambda_{\dot\alpha}. \qquad (2.4)$$

Note that λ_α and $\tilde\lambda_{\dot\alpha}$ are sometimes called "holomorphic" and "anti-holomorphic" spinors because of their transformation properties under the Lorentz group. Bear in mind that there is no connection here to any underlying fermionic states—these spinors are *not* anti-commuting Grassmann numbers, but rather plain old complex numbers.

For real momenta, $p_{\alpha\dot\alpha}$ is Hermitian, implying the reality condition,

$$\tilde\lambda_{\dot\alpha} = \pm(\lambda^*)_{\dot\alpha}, \qquad (2.5)$$

for positive and negative energy states, respectively. When the momentum is complex, there is no relation between λ_α and $\tilde{\lambda}_{\dot\alpha}$ and they are independent. However, in this case $p_{\alpha\dot\alpha}$ still satisfies Eq. (2.3) so these spinors still describe a massless particle. Strictly speaking, since the external momenta in physical processes are real, we should maintain this reality condition throughout our calculations. However, as we will soon see, much of our power will flow from analytic continuation to complex momenta while treating the physical S-matrix as a restriction of this more general object.

We are now ready to introduce the Lorentz invariant building blocks of the spinor helicity formalism. Given two particles i and j, we define

$$\langle ij \rangle = \lambda_{i\alpha}\lambda_{j\beta}\epsilon^{\alpha\beta} \tag{2.6}$$

$$[ij] = \tilde{\lambda}_{i\dot\alpha}\tilde{\lambda}_{j\dot\beta}\epsilon^{\dot\alpha\dot\beta}, \tag{2.7}$$

which are commonly referred to as "angle" and "square" brackets. Any function of four-dimensional kinematic data can be written exclusively in terms of these objects. For example, angle and square brackets come together to form the familiar Mandelstam invariants,

$$s_{ij} = (p_i + p_j)^2 = 2p_ip_j = \langle ij \rangle[ij], \tag{2.8}$$

and likewise for higher-point kinematic invariants.

Because spinors are simple objects, they are subject to relatively few algebraic manipulations. One important property is antisymmetry, $\langle ij \rangle = -\langle ji \rangle$ and $[ij] = -[ji]$, which in turn implies that $\langle ii \rangle = [ii] = 0$. Moreover, since each spinor is two-dimensional, one can always write a spinor as a linear combination of two linearly independent spinors,

$$\langle ij \rangle\lambda_k + \langle ki \rangle\lambda_j + \langle jk \rangle\lambda_i = 0, \tag{2.9}$$

which is known as the Schouten identity.

In point of fact, spinor helicity variables are nothing more than an algebraic reshuffling of the external kinematic data. Such a manipulation would not be particularly advantageous were it not for the fact that scattering amplitudes enjoy an *immense* reduction in complexity when translated into these variables. The avatar of this simplification is the S-matrix for gluon scattering. Depending on the level of sadism in your introductory QFT course, it is quite likely that you have computed the four-gluon amplitude and possibly even the five-gluon amplitude but not higher. The amplitude for six-particle

scattering was completed in a brute-tour-de-force Feynman diagram calculation of Parke and Taylor [19]. The latter involves two hundred and twenty diagrams, spelled out in several pages of nested variables and tables of numerical coefficients. Shortly after, Parke and Taylor brilliantly realized that for a particular "maximally helicity violating" (MHV) helicity configuration, *i.e.* two negative helicity gluons with the rest positive, this mountain of algebra telescopes into an expression built from simple color structures multiplying *monomial* expressions of the form

$$A(\cdots i^- \cdots j^- \cdots) = \frac{\langle ij \rangle^4}{\langle 12 \rangle \langle 23 \rangle \cdots \langle n1 \rangle}, \qquad (2.10)$$

which is the celebrated Parke-Taylor formula [20].

So what is the moral of this particular story? First, we see with brutal clarity that almost all of the terms in the Feynman diagram expansion are nothing more than a *complicated rewriting of zero*. Second, new features of the amplitude become manifest: in particular, the Parke-Taylor formula is purely a function angle brackets, revealing an underlying holomorphic structure of MHV amplitudes. As observed long ago by Nair [21], this property led in part to the twistor-string description of gluon scattering [22] as well as more recent developments connecting soft gluon radiation to an underlying conformal field theory [23, 24].

2.2   Little Group

Spinor helicity variables have a certain aesthetic elegance. But why are they practically useful? As is so often the case, their utility flows from the fact that they linearly realize the symmetries of the system. These symmetries are Lorentz invariance and the so-called little group, which we now discuss.

The little group is comprised of the subset of Lorentz transformations which leave the momentum p_μ of a particle invariant. According to the Wigner classification [1, 2], a particle of momentum p_μ is then described by an irreducible representation of the little group. For example, consider the momentum vector of a massless particle in four dimensions,

$$p_\mu = (p_0, 0, 0, p_0), \qquad (2.11)$$

shown here in a certain frame. The transformations which leave p_μ invariant form the $ISO(2)$ group of translations and rotations in the plane transverse

to the trajectory of the particle. The finite dimensional representations of this group are eigenvectors of the rotation subgroup. The corresponding eigenvalues label the helicities of physical states.

Under the little group, the spinor helicity variables λ_i and $\tilde{\lambda}_i$ transform so as to leave p_i is unchanged, so

$$\lambda_i \to t_i \lambda_i \qquad \text{and} \qquad \tilde{\lambda}_i \to t_i^{-1} \tilde{\lambda}_i. \tag{2.12}$$

Note that for real momenta, the spinors satisfy a reality condition so for physical kinematics t_i is restricted to be a pure phase.

Of what use is the little group if it leaves all momenta p_i invariant? For interacting scalars, the answer is *of no use at all*, since these particles are singlets of the little group and the amplitude is solely a function of p_i. On the other hand, for particles with spin it is necessary to introduce nontrivial representations of the little group which carry helicity quantum numbers. In terms of conventional Feynman diagrams, this little group covariance enters through an additional set of kinematic objects we have thus far ignored: the polarizations.

In terms of spinor helicity variables, polarization vectors take the form

$$e^+_{\alpha\dot\alpha} = \frac{\eta_\alpha \tilde{\lambda}_{\dot\alpha}}{\langle \eta\lambda \rangle} \qquad \text{and} \qquad e^-_{\alpha\dot\alpha} = \frac{\lambda_\alpha \tilde{\eta}_{\dot\alpha}}{[\tilde{\lambda}\tilde{\eta}]}, \tag{2.13}$$

where the $\pm$ superscripts label helicity. By construction, these polarizations are transverse to the momenta, so

$$p^{\alpha\dot\alpha} e^+_{\alpha\dot\alpha} \propto [\tilde{\lambda}\tilde{\lambda}] = 0 \qquad \text{and} \qquad p^{\alpha\dot\alpha} e^-_{\alpha\dot\alpha} \propto \langle \lambda\lambda \rangle = 0. \tag{2.14}$$

Note the appearance of "reference spinors", η and $\tilde{\eta}$, which are linearly independent of λ and $\tilde{\lambda}$ but otherwise arbitrary. In fact, one can even assign *different* reference spinors for each external particle. How is this possible? If we change η and $\tilde{\eta}$, should not the answer change? To see why this yields no contradiction, consider substituting the reference spinor η with a different but still arbitrary reference spinor η',

$$\eta_\alpha \to \eta'_\alpha = a\eta_\alpha + b\lambda_\alpha, \tag{2.15}$$

where we have expanded η' as a linear combination of the original reference spinor η and the momentum spinor λ. The corresponding polarization vector

is similarly shifted according to

$$e^+_{\alpha\dot\alpha} \to e^{+\prime}_{\alpha\dot\alpha} = e^+_{\alpha\dot\alpha} + \left(\frac{b}{a\langle\eta\lambda\rangle}\right) p_{\alpha\dot\alpha}, \tag{2.16}$$

which is literally a gauge transformation. So the polarizations *do* change depending on the choice of reference but in a way that leaves the amplitude invariant by the Ward identity.

Under the little group, the polarization vectors transform as

$$e^+_{\alpha\dot\alpha} \to t^{-2} e^+_{\alpha\dot\alpha} \qquad \text{and} \qquad e^-_{\alpha\dot\alpha} \to t^2 e^-_{\alpha\dot\alpha}, \tag{2.17}$$

so the little group covariance of each polarization encodes its associated helicity. A scattering amplitude with multiple particles will then be multilinear in the corresponding polarizations, so

$$A(1^{h_1}\cdots n^{h_n}) = e^{h_1}_{\mu_1}\ldots e^{h_n}_{\mu_n} A^{\mu_1\cdots\mu_n}, \tag{2.18}$$

where $h_i = \pm 1$ labels the helicity of each leg. Hence, the scattering amplitude is little group covariant with weight

$$A(1^{h_1}\cdots n^{h_n}) \to \prod_i t_i^{-2h_i} A(1^{h_1}\cdots n^{h_n}). \tag{2.19}$$

At first glance it may seem peculiar that the scattering amplitude is covariant rather than invariant under the little group. After all, all physical observables are invariant. Nevertheless, there is no contradiction because on real kinematics the little group parameters t_i reduce to pure phases which trivially cancel in the matrix element squared.

The tensorial object on the right-hand side of Eq. (2.18), $A^{\mu_1\cdots\mu_n}$, is the travesty computed from Feynman diagrams in the standard approach. Meanwhile, the object on left-hand side, $A(1^{h_1}\cdots n^{h_n})$, labels "helicity amplitudes," which are a set of distinct functions corresponding to each helicity configuration. These objects are *true* scattering amplitudes in the sense that they are fully gauge invariant functions like the Parke-Taylor formula in Eq. (2.10). The aim of the amplitudes program is to sidestep $A^{\mu_1\cdots\mu_n}$ altogether by bootstrapping $A(1^{h_1}\cdots n^{h_n})$ directly from the physical principles outlined previously.

Reference spinors can be tremendously useful for explicit calculations. As a simple example, consider the amplitude for all plus helicity gluons,

$A(1^+2^+\cdots n^+)$. First, note that every term in the Feynman diagram expansion appears with at least one factor of $e_i^+e_j^+$. This is because the amplitude has mass dimension $4-n$ and at most $n-3$ propagators. Thus, the numerator of every Feynman diagram has $n-2$ powers of momenta or fewer, implying that at least two polarization vectors are contracted via $e_i^+e_j^+$. However, by choosing the *same* reference spinor, η, for all polarizations, we set $e_i^+e_j^+ = 0$, establishing that $A(1^+2^+\cdots n^+) = 0$. A similar argument applies to the all but one plus helicity gluon amplitude, $A(1^-2^+\cdots n^+)$. Again, every term enters with at least one factor of $e_i^+e_j^+$, so choosing $\eta = \lambda_1$ for every reference sets $e_i^+e_j^+ = e_1^-e_i^+ = 0$ so $A(1^-2^+\cdots n^+) = 0$ as well. Without even doing a calculation, we have proven that the leading nontrivial tree-level gluon scattering amplitudes is MHV.

3   Bootstrapping Amplitudes

The goal of this section is to systematically enumerate all possible Lorentz invariant interactions among massless particles in four dimensions. In accordance with our stated philosophy these results will be derived without reference to an action. This concept of an "amplitudes bootstrap" closely follows the seminal work of Benincasa and Cachazo [25], though similar treatments can be found in the existing review literature [26, 27].

3.1   Three-Particle Amplitudes

The leading nontrivial contribution to the S-matrix is the three-particle amplitude. According to momentum conservation,

$$p_1 + p_2 + p_3 = 0 \qquad \Rightarrow \qquad \begin{array}{l} (p_1 + p_2)^2 = \langle 12\rangle[12] = p_3^2 = 0 \\ (p_2 + p_3)^2 = \langle 23\rangle[23] = p_1^2 = 0 \\ (p_3 + p_1)^2 = \langle 31\rangle[31] = p_2^2 = 0 \end{array} \qquad (3.1)$$

From the top line, we deduce that if $\langle 12\rangle \neq 0$, then $[12] = 0$. Furthermore, $\langle 12\rangle[23] = -\langle 11\rangle[13] - \langle 13\rangle[33] = 0$, so $[23] = 0$. Repeating this procedure cyclically, we find that $[12] = [23] = [31] = 0$. If, on the other hand, we assume $[12] \neq 0$, then $\langle 12\rangle = \langle 23\rangle = \langle 31\rangle = 0$. In summary, the three-particle amplitude only has support on two possible kinematic configurations: holomorphic, corresponding to all vanishing square brackets,

$$[12] = [23] = [31] = 0 \qquad \Rightarrow \qquad \tilde{\lambda}_1 \propto \tilde{\lambda}_2 \propto \tilde{\lambda}_3, \qquad (3.2)$$

and anti-holomorphic, corresponding to all vanishing angle brackets,

$$\langle 12 \rangle = \langle 23 \rangle = \langle 31 \rangle = 0 \qquad \Rightarrow \qquad \lambda_1 \propto \lambda_2 \propto \lambda_3. \tag{3.3}$$

Both kinematic configurations imply that $p_1 p_2 = p_2 p_3 = p_3 p_1 = 0$ and require complex momenta so that λ_i and $\tilde{\lambda}_i$ are independent variables.

Without loss of generality, the three-particle amplitude in the holomorphic kinematic configuration takes the form

$$A(1^{h_1} 2^{h_2} 3^{h_3}) = \langle 12 \rangle^{n_3} \langle 23 \rangle^{n_1} \langle 31 \rangle^{n_2}. \tag{3.4}$$

Imposing the criterion of little group covariance in Eq. (2.19), we solve for the exponents n_i in terms of the helicities h_i of the external particles,

$$
\begin{aligned}
-2h_1 &= n_2 + n_3 & n_1 &= h_1 - h_2 - h_3 \\
-2h_2 &= n_3 + n_1 \quad \Rightarrow \quad & n_2 &= h_2 - h_3 - h_1 \\
-2h_3 &= n_1 + n_2 & n_3 &= h_3 - h_1 - h_2
\end{aligned}
\tag{3.5}
$$

Note that the exponents are integers because all helicities h_i are integers or half-integers and fermions always appear in pairs.

The mass dimension of the three-particle amplitude cannot be negative, as this would require nonlocality due to inverse powers of derivatives in the three-particle vertex. Hence, the assumption of locality implies that the three-particle amplitude has nonnegative mass dimension, so

$$[A(1^{h_1} 2^{h_2} 3^{h_3})] = n_1 + n_2 + n_3 = -(h_1 + h_2 + h_3) = -h \geq 0, \tag{3.6}$$

where $[\ldots]$ denotes the mass dimension accumulated by powers of momenta, ignoring the intrinsic dimensionality of coupling constants. From Eq. (3.6) we see that holomorphic kinematics applies for $h \leq 0$ while anti-holomorphic kinematics applies for $h \geq 0$. In summary, we have derived a general formula for the three-particle amplitude of massless particles in four dimensions,

$$A(1^{h_1} 2^{h_2} 3^{h_3}) = \begin{cases} \langle 12 \rangle^{h_3 - h_1 - h_2} \langle 23 \rangle^{h_1 - h_2 - h_3} \langle 31 \rangle^{h_2 - h_3 - h_1}, & h \leq 0 \\ [12]^{h_1 + h_2 - h_3} [23]^{h_2 + h_3 - h_1} [31]^{h_3 + h_1 - h_2}, & h \geq 0 \end{cases} \tag{3.7}$$

which like the tuxedo t-shirt is both practical and elegant.

Since this formula was derived purely from symmetry and dimensional analysis, it holds *nonperturbatively*. For example, loops which would naively generate logarithms or more complicated functional objects like $\log(p_1 p_2)$ are all singular or vanishing on three-particle kinematics.

In the subsequent sections, we will study Eq. (3.7) for the familiar cases of scalars, vectors, and tensors.

3.1.1 Scalars

For identical scalars, all helicities $h_i = 0$ are vanishing and Eq. (3.7) implies that the three-particle amplitude is a constant,

$$A(123) = \omega, \tag{3.8}$$

corresponding to a single scalar self-interacting through a cubic potential term. A similar result applies to multiple scalars,

$$A(1_a 2_b 3_c) = \omega_{abc}, \tag{3.9}$$

where the subscripts run over flavors. Since the external states are bosons, the coupling constant ω_{abc} is symmetric in its indices.

We have argued that Eq. (3.9) is the only possible on-shell three-particle amplitude of massless scalars. But what about derivatively coupled scalars, *e.g.* interacting through $\phi(\partial\phi)^2$? In fact, these amplitudes vanish on-shell because they can only depend on products of momenta and these necessarily vanish: $p_1^2 = p_2^2 = p_3^2 = p_1 p_2 = p_2 p_3 = p_3 p_1 = 0$. At the level of the action, this triviality can be made manifest by applying a field redefinition that simply eliminates these cubic derivative interactions altogether.

3.1.2 Vectors

For the case of identical vectors the helicities are $h_i = \pm 1$. This implies that the exponents n_i in Eq. (3.4) are odd integers. Consequently, the three-particle amplitude is *odd* under the exchange of any two external states—even though they are bosons. Self-consistency then demands that the three-particle amplitude must identically *vanish*. This conclusion is in perfect agreement with the textbooks: the three-particle amplitude of photons is indeed zero by charge conjugation symmetry.

While this argument forbids cubic self-interactions of a single vector, it still allows for interactions among multiple species—also known as colors—of vectors. Introducing a color index to each vector, we obtain

$$A(1_a^- 2_b^- 3_c^+) = f_{abc} \frac{\langle 12 \rangle^3}{\langle 13 \rangle \langle 32 \rangle} \qquad \text{and} \qquad A(1_a^+ 2_b^+ 3_c^-) = f_{abc} \frac{[12]^3}{[13][32]}. \tag{3.10}$$

Here f_{abc} has to be fully antisymmetric in its indices so that the amplitude is *even* under exchange of bosons. These expressions exactly agree with the

more complicated expressions obtained from Feynman diagrams,

$$A(1_a 2_b 3_c) = f_{abc}(e_1 e_2)(p_1 e_3 - p_2 e_3) + \text{cyclic perm}, \qquad (3.11)$$

after inserting the expressions for the polarization vectors in Eq. (2.13).

For the all minus and all plus helicity combinations, Eq. (3.7) reduces to

$$A(1_a^- 2_b^- 3_c^-) = f_{abc}\langle 12\rangle\langle 23\rangle\langle 31\rangle, \qquad A(1_a^+ 2_b^+ 3_c^+) = f_{abc}[12][23][31], \quad (3.12)$$

where f_{abc} is again antisymmetric so the amplitude is even under exchange of bosons. By counting powers of momentum, it is clear that Eq. (3.12) originates from a three-derivative interaction among gluons. This arises from a *nonrenormalizable* higher dimension operator, $f_{abc}F_\mu^{a\nu}F_\nu^{b\rho}F_\rho^{c\mu}$, which is generated from loops of heavy colored particles.

3.1.3   Tensors

For the case of identical tensors the helicities are $h_i = \pm 2$. From Eq. (3.4), we see that the exponents n_i are even, so the amplitude is invariant under the exchange of bosons. The resulting three-particle amplitude is

$$A(1^{--}2^{--}3^{++}) = \frac{\langle 12\rangle^6}{\langle 13\rangle^2\langle 32\rangle^2}, \qquad A(1^{++}2^{++}3^{--}) = \frac{[12]^6}{[13]^2[32]^2}, \quad (3.13)$$

corresponding to the scattering of three gravitons. As before, we can compare this expression to the equivalent Feynman diagram expression. Even with the aid of a simplified action for graviton perturbations [28], we obtain

$$A(123) = -\frac{1}{2}(e_1 e_2)^2 (p_1 e_3)(p_2 e_3) + (e_1 e_2)(e_2 e_3)(p_1 e_3)(p_2 e_1) + \text{perm}, \tag{3.14}$$

which is significantly more complicated then Eq. (3.13). Note that here and for the remainder of these lectures, all tensor polarizations will be written as products of vector polarizations.

Last but not least, consider the amplitude for tensors scattering in the all minus and all plus configurations,

$$A(1^{--}2^{--}3^{--}) = \langle 12\rangle^2\langle 23\rangle^2\langle 31\rangle^2, \qquad A(1^{++}2^{++}3^{++}) = [12]^2[23]^2[31]^2. \tag{3.15}$$

Again counting powers of momentum, we see that these amplitudes are generated by a six-derivative interaction among gravitons induced by the curvature-cubed operator, $R_{\mu\nu}{}^{\rho\sigma}R_{\rho\sigma}{}^{\alpha\beta}R_{\alpha\beta}{}^{\mu\nu}$.

It is peculiar that we have jumped so quickly to *cubic* curvature invariants. After all, what happened to the *quadratic* curvature invariants, R^2, $R_{\mu\nu}R^{\mu\nu}$, and $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$? By explicit calculation, one can verify that these operators produce demonstrably nontrivial Feynman vertices. Have we somehow missed the four-derivative contributions to the amplitude?

On the contrary, the above calculations are perfectly sound. Moreover, they actually foreshadow one of the great surprises in the history of quantum gravity. Here we refer to the pioneering work of 't Hooft and Veltman [29], who first calculated the one-loop divergences of four-dimensional pure gravity. By naive power counting, one-loop graviton diagrams should produce divergent contributions corresponding to local curvature-squared counterterms like R^2, $R_{\mu\nu}R^{\mu\nu}$ and $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$. There is, however, a huge caveat. None of these operators actually contribute to on-shell scattering amplitudes in four dimensions! You can verify this explicitly—the off-shell Feynman diagrams are complicated and non-trivial but their effects evaporate on-shell.

To understand this at the level of the action, rewrite $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$ in terms of the Gauss-Bonnet combination, $R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$. In four dimensions this is a total derivative which decouples from perturbative scattering. Exploiting the freedom of field redefinitions, then plug the leading order equations of motion—that is, Einstein's field equations—into the next-to-leading order corrections. This sets $R = R_{\mu\nu} = 0$ into the curvature-squared operators, eliminating them altogether.

The upshot here is twofold: in four dimensions, curvature-squared operators do not contribute to the tree-level graviton S-matrix, and pure gravity is *finite* at one-loop. This second fact briefly prompted optimism that pure gravity might be finite at higher orders—a flighting dream soon crushed by a two loop divergence [30, 31, 32, 33] corresponding to the $R_{\mu\nu}{}^{\rho\sigma}R_{\rho\sigma}{}^{\alpha\beta}R_{\alpha\beta}{}^{\mu\nu}$ counterterm.

Nowadays, the divergence structure of supergravity theories remains an active field of study. Explicit calculations have exposed unexpected cancellations not anticipated by counterterm arguments [34, 35, 36, 37, 38, 39], suggesting the enticing possibility that $\mathcal{N} = 8$ supergravity might actually be ultraviolet finite to *all orders*. These developments call to mind the time-worn adage, "all that is not forbidden is compulsory." Though true at a buffet, such wisdom must be applied with care to $\mathcal{N} = 8$ supergravity.

3.2 Four-Particle Amplitudes

That the three-particle amplitude is wholly fixed by symmetry is perhaps unsurprising. After all, these objects are generated by pure contact vertices and do not suffer from the complications of exchanged particles. For the same reason, extending our analysis to four-particle amplitudes will require a new ingredient: locality.

Locality is encoded in the singularity structure of an amplitude. In particular, consider when the external kinematics are carefully tuned so that an intermediate particle is on-shell. The tree-level amplitude exhibits a $1/p^2$ singularity signalling the propagation of the intermediate state over a macroscopic physical distance in spacetime. In this factorization limit the amplitude decomposes into a product of subamplitudes corresponding to two distinct processes occurring at disparate points.

This factorization property places stringent conditions on the structure of the amplitude. For example, in the four-particle amplitude, simple poles like $1/s$ can arise, but not poles like $1/s^2$. For the tree-level four-particle amplitude, this implies that

$$\lim_{s \to 0} s \, A_4 = A_3 \, A_3, \tag{3.16}$$

which is a highly nontrivial relationship linking higher-point and lower-point kinematics. As a trivial corollary, the mass dimensions of the four-particle and three-particle amplitudes are related by

$$[A_4] = 2[A_3] - 2, \tag{3.17}$$

where as before, our definition of mass dimension tabulates contributions from kinematic factors but not couplings.

In the remainder of this section, we derive the four-particle tree amplitudes for massless particles purely from factorization and our results for three-particle amplitudes.

3.2.1 Scalars

As we saw earlier, the three-particle amplitude has positive mass dimension, $[A_3] \geq 0$, since the associated vertex should be local. By dimensional analysis, Eq. (3.17) then implies that

$$[A_3] \geq 0 \qquad \Rightarrow \qquad [A_4] \geq -2, \tag{3.18}$$

so the mass dimension of the four-particle amplitude of scalars is bounded from below. Furthermore, the tree amplitude should be a permutation invariant function of s, t, and u with only simple poles. Enumerating all possible such functions, we obtain

$$\phi^3 \quad : \quad A_4 = s^{-1} + t^{-1} + u^{-1} \tag{3.19}$$

$$\phi^4 \quad : \quad A_4 = 1 \tag{3.20}$$

$$(\partial\phi)^2\phi^2 \quad : \quad A_4 = s + t + u = 0 \tag{3.21}$$

$$(\partial\phi)^4 \quad : \quad A_4 = s^2 + t^2 + u^2 \tag{3.22}$$

$$(\partial\partial\phi)^2(\partial\phi)^2 \quad : \quad A_4 = s^3 + t^3 + u^3, \tag{3.23}$$

and so on. Note that the ϕ^3 amplitude has singularities in every factorization channel. In contrast, all the derivatively coupled scalar amplitudes are regular, consistent with the fact that their associated on-shell three-particle amplitudes are zero according to our previous arguments.

3.2.2   Vectors

Next, consider the four-particle tree amplitude of massless vectors. From our earlier analysis we saw that $[A_3] = 1$, so Eq. (3.17) implies that

$$[A_3] = 1 \qquad \Rightarrow \qquad [A_4] = 2[A_3] - 2 = 0, \tag{3.24}$$

and the amplitude is dimensionless.

By incorporating the constraints of little group covariance and dimensional analysis, it is possible to build a restrictive ansatz for the four-particle amplitude. To do so, we exploit that amplitude transforms as

$$A(1_a^- 2_b^- 3_c^+ 4_d^+) \rightarrow t_1^2 t_2^2 t_3^{-2} t_4^{-2} A(1_a^- 2_b^- 3_c^+ 4_d^+), \tag{3.25}$$

under the little group. Factoring out the full little group weight of the amplitude, we consider the general ansatz,

$$A(1_a^- 2_b^- 3_c^+ 4_d^+) = \langle 12 \rangle^2 [34]^2 F(s, t, u), \tag{3.26}$$

where we have defined

$$F(s, t, u) = \frac{c_{st}}{st} + \frac{c_{tu}}{tu} + \frac{c_{us}}{us}, \tag{3.27}$$

which is the most general little group invariant function of proper mass dimension and with only simple poles. By dimensional analysis, the coefficients c_{st}, c_{tu}, c_{us} are dimensionless constants.

Demanding factorization on the s-channel singularity, we find that

$$
\begin{aligned}
\lim_{s \to 0} s\, A(1_a^- 2_b^- 3_c^+ 4_d^+) &= \langle 12 \rangle^2 [34]^2 \frac{1}{t} \left(c_{st} - c_{us} \right) \\
&= \sum_{h=\pm} \sum_e A(1_a^- 2_b^- P_e^h) A(3_c^+ 4_d^+ P_e^{-h}) = \sum_e A(1_a^- 2_b^- P_e^+) A(3_c^+ 4_d^+ P_e^-) \\
&= \sum_e f_{abe} f_{cde} \frac{\langle 12 \rangle^3}{\langle 1P \rangle \langle P2 \rangle} \frac{[34]^3}{[3P][P4]} = \sum_e f_{abe} f_{cde} \langle 12 \rangle^2 [34]^2 \frac{1}{t},
\end{aligned}
\tag{3.28}
$$

where we have defined $P = -(p_1 + p_2) = (p_3 + p_4)$ and set $t = -u$ on the factorization channel. In the last line we have used that $\langle 1P \rangle [P4] = -\langle 12 \rangle [24]$ and $\langle P2 \rangle [3P] = \langle 42 \rangle [34]$.

Comparing the first and last lines of Eq. (3.28), we obtain an equation for $c_{st} - c_{us}$ in terms of the structure constants. Repeating this exercise for the t- and u-channels we obtain a triplet of equations,

$$
c_{st} - c_{us} = \sum_e f_{abe} f_{cde}
\tag{3.29}
$$

$$
c_{tu} - c_{st} = \sum_e f_{bce} f_{ade}
\tag{3.30}
$$

$$
c_{us} - c_{tu} = \sum_e f_{cae} f_{bde},
\tag{3.31}
$$

whose sum is an equation with a very old name,

$$
\sum_e f_{abe} f_{cde} + f_{bce} f_{ade} + f_{cae} f_{bde} = 0,
\tag{3.32}
$$

i.e. the Jacobi identity. In the textbook approach to YM theory, the Jacobi identity is a byproduct of an underlying Lie algebra—nothing more than an algebraic identity, $\mathrm{tr}([T_a, T_b][T_c, T_d]) + \mathrm{tr}([T_b, T_c][T_a, T_d]) + \mathrm{tr}([T_c, T_a][T_b, T_d]) = 0$. Here the Jacobi identity is a *consistency condition* arising from factorization of the four-particle amplitude!

3.2.3 Tensors

Last but not least is the four-particle tree amplitude of massless tensors. From our earlier analysis we learned that the mass dimension of the three-

particle graviton amplitude is $[A_3] = 2$, so Eq. (3.17) implies that

$$[A_3] = 2 \quad \Rightarrow \quad [A_4] = 2[A_3] - 2 = 2. \tag{3.33}$$

Like before, we construct a general ansatz for the four-particle amplitude consistent with little group covariance and dimensional analysis,

$$A(1^{--}2^{--}3^{++}4^{++}) = \langle 12 \rangle^4 [34]^4 F(s, t, u), \tag{3.34}$$

where $F(s, t, u)$ is the unique little group and permutation invariant object of the correct mass dimension with only simple poles,

$$F(s, t, u) \propto \frac{1}{stu}. \tag{3.35}$$

The above formulas agree with the known expression for the four-particle graviton amplitude. Note that we have arrived at this result without even imposing factorization.

3.3   Diverse Dimensions

Thus far we have restricted to four dimensions in order to reap the benefits of the spinor helicity formalism. As one might expect, however, an analogous version of the bootstrap can be implemented in d dimensions. Starting from an ansatz constructed from d-dimensional kinematic invariants, one imposes consistency conditions like gauge invariance and factorization in order to excise the physical scattering amplitude.

Admittedly, this approach is an unabashed retreat from our stated principles. After all, we had vowed earlier to banish the notion of gauge invariance from all future discussions. Nevertheless, the following exercise gives a flavor of the mechanics of the amplitudes bootstrap in diverse dimensions.

To begin, consider the three-particle amplitude. The complete basis of independent Mandelstam invariants for three particles is[3]

$$\{p_1 e_2, p_1 e_3, p_2 e_1, e_1 e_2, e_1 e_3, e_2 e_3\}. \tag{3.36}$$

For the case of gluons, the amplitude is linear in e_1, e_2, and e_3, so the most general ansatz with the power counting of YM theory is

$$A(1_a 2_b 3_c) = c_1(e_1 e_3)(p_1 e_2) + c_2(e_1 e_2)(p_1 e_3) + c_3(e_2 e_3)(p_2 e_1). \tag{3.37}$$

---

[3] See Appendix B for a derivation of this basis for any number of particles.

By applying the Ward identity to each external leg, one can verify that the only combination of terms that satisfies the Ward identities is proportional to $(c_1, c_2, c_3) \propto (-1, 1, 1)$, which agrees with the expression from Feynman diagrams as well as the result in Eq. (3.10).

For gravitons, the amplitude is instead quadratic in e_1, e_2, and e_3, so the most general ansatz with the power counting of gravity is

$$
\begin{aligned}
A(123) = \quad & d_1 \, (e_1 e_3)^2 \, (p_1 e_2)^2 + d_2 \, (e_1 e_2) \, (e_1 e_3) \, (p_1 e_3) \, (p_1 e_2) \\
& + d_3 \, (e_1 e_2)^2 \, (p_1 e_3)^2 + d_4 \, (e_1 e_3) \, (e_2 e_3) \, (p_2 e_1) \, (p_1 e_2) \\
& + d_5 \, (e_1 e_2) \, (e_2 e_3) \, (p_1 e_3) \, (p_2 e_1) + d_6 \, (e_2 e_3)^2 \, (p_2 e_1)^2.
\end{aligned} \tag{3.38}
$$

Similarly, one can check that the only diffeomorphism invariant combination is proportional to $(d_1, d_2, d_3, d_4, d_5, d_6) \propto (1, -2, 1, -2, 2, 1)$, which again matches Feynman diagrams and our expression derived in Eq. (3.13).

It is straightforward to extend these arguments to higher point by including factorization as an additional input. We will not go through this exercise here. However, it has been shown [40, 41, 42] that an even milder set assumptions is actually sufficient to derive the S-matrix. In fact, one can generate the full tree-level S-matrices of YM and gravity using only gauge invariance and certain weak assumptions about the singularity structure.

Note that the graviton three-particle amplitude is *literally* the square of the gluon three-particle amplitude in any dimension. At face value this is absolutely miraculous, though anticipated by the four-dimensional expressions derived in Eq. (3.10) and Eq. (3.13). Nevertheless, given our bootstrapping approach it should be obvious that this relationship was nearly *inevitable*. In particular, since the three-particle amplitude of gluons is a gauge invariant function, so too is its square. This squared object matches the power counting of gravity, so it is automatically a viable candidate function for the three-particle amplitude of gravitons. Hence, the only real surprise is that there are no other gauge invariant objects *beyond this* and the graviton amplitude is unique.

4   Recursion Relations

The previous sections describe a first principles derivation of all possible three- and four-particle scattering amplitudes. While these methods straightforwardly generalize to higher-point, there is thankfully a less artisanal approach. In particular, on-shell recursion *automates* this factorization proce-

dure by rewriting on-shell amplitudes in terms of lower-point on-shell amplitudes. When the entire tree-level S-matrix is defined recursively from a finite number of seed amplitudes, the corresponding theory is said to be *on-shell constructible*. Pioneered in the seminal work of Britto, Cachazo, Feng, and Witten [43, 44], on-shell recursion has since blossomed into a veritable cottage industry.

At first glance, on-shell recursion relations share a superficial similarity with Feynman diagrams. After all, every Feynman diagram is computed as a sum over products of subdiagrams. In fact, Berends and Giele long ago derived compact *off-shell* recursion relations [45] which efficiently repackage the Feynman diagrammatic expansion. Despite this, on-shell recursion offers the important conceptual benefit of a purely *on-shell* definition of the S-matrix. In this picture, the off-shell "bulk" spacetime associated with the action formalism is demonstrably unnecessary.

In what follows, we derive on-shell recursion relations for scattering amplitudes in a broad range of QFTs. Our discussion will be light on explicit examples, focusing more on the broad implications for QFT in general. For the avid reader seeking the gory details, we suggest the excellent treatments of on-shell recursion in the review literature [26, 27].

4.1   Momentum Shifts

On-shell recursion is a systematic procedure for relating an amplitude to its values at singular kinematics. In order to probe these kinematic configurations we define a momentum shift, which is a one-parameter deformation of the external momenta engineered to sample various kinematic limits. Here we will consider shifts of the form

$$p_i \to p_i(z) = p_i + zq_i, \tag{4.1}$$

where z is a complex number parameterizing deviations from the original kinematic configuration. Note that some of the q_i may be zero if there are unshifted external legs. The shifted momenta in turn induce a deformation of the amplitude,

$$A \to A(z). \tag{4.2}$$

It is crucial that the momentum shift preserve on-shell kinematics for *all z*, so the deformed amplitude $A(z)$ is still perfectly physical despite its gruesome

moniker. Hence, the momentum shifts are restricted by total momentum conservation,

$$\sum_i p_i(z) = 0 \qquad \Rightarrow \qquad \sum_i q_i = 0, \tag{4.3}$$

together with the on-shell conditions,

$$p_i(z)^2 = 0 \qquad \Rightarrow \qquad q_i^2 = q_i p_i = 0. \tag{4.4}$$

Eq. (4.3) and Eq. (4.4) should be interpreted as conditions on q_i. These equations are generally easy to satisfy.

4.2 Factorization from Analyticity

At last we are ready to derive a general formula for on-shell recursion. To begin, we exploit the trivial fact that the original amplitude, $A(0)$, is obtained from the residue of $A(z)$ at $z = 0$,

$$A(0) = \frac{1}{2\pi i} \oint_{z=0} dz \, \frac{A(z)}{z} = -\sum_I \underset{z=z_I}{\text{Res}} \left[\frac{A(z)}{z} \right] + B_\infty, \tag{4.5}$$

where we have used Cauchy's theorem to express the residue at $z = 0$ in terms of the residues at all other poles, z_I, plus a boundary contribution, B_∞. Crucially, this manipulation is only permitted if $A(z)$ is a *rational* function of the kinematics, *i.e.* the amplitude is tree-level. Also, note that $B_\infty = 0$ whenever $A(z)$ vanishes at large z.

Each pole in z_I labels a kinematic singularity of the amplitude. On physical grounds, this *necessarily* coincides with a physical factorization channel. To see this explicitly, consider a subset I of external particle carrying total momentum,

$$P_I = \sum_{i \in I} p_i \qquad \Rightarrow \qquad P_I(z) = \sum_{i \in I} p_i(z), \tag{4.6}$$

after the momentum shift. Here the shifted total momentum is

$$P_I(z) = P_I + z Q_I, \tag{4.7}$$

where $Q_I = \sum_{i \in I} q_i$. For general kinematics, $P_I^2 \neq 0$ and so P_I is off-shell. However, the deformed momentum $P_I(z)$ *is* on-shell when

$$0 = P_I^2(z_I) = P_I^2 + 2z_I Q_I P_I \qquad \Rightarrow \qquad z_I = -\frac{P_I^2}{2Q_I P_I}. \tag{4.8}$$

For simplicity we have assumed that the $q_i q_j = 0$ so that $Q_I^2 = 0$ for any subset of legs I. Strictly speaking this is not mandatory but the calculation simplifies substantially when $P_I^2(z)$ is a linear function. For the more general case of $Q_I^2 \neq 0$, $P_I^2(z)$ is quadratic in z and contributions from this factorization channel will be split between two singular points corresponding to the roots of this polynomial [46].

At the point z_I, the amplitude should factorize into a product of lower-point amplitudes,

$$\lim_{z \to z_I} P_I^2(z) A(z) = A_L(z_I) A_R(z_I), \tag{4.9}$$

which implies that the associated residue is

$$\operatorname*{Res}_{z=z_I} \left[\frac{A(z)}{z} \right] = \frac{1}{z_I} \left(\frac{dP_I^2(z_I)}{dz_I} \right)^{-1} \times \lim_{z \to z_I} \left[P_I^2(z) A(z) \right] \tag{4.10}$$

$$= -\frac{1}{P_I^2} \times A_L(z_I) A_R(z_I). \tag{4.11}$$

Plugging back into Eq. (4.5), we obtain our final expression

$$A(z=0) = \sum_I \frac{1}{P_I^2} A_L(z_I) A_R(z_I) + B_\infty, \tag{4.12}$$

which is a general formula for on-shell recursion. Here I runs over all singularities z_I in the amplitude.

While each z_I pole corresponds to a factorization channel, the converse may not be true. In fact, the only factorization channels which enter into Eq. (4.12) are those that have shifted momenta on *both* sides of the channel. If instead all the shifted momenta were on *one* side of the channel, then $P_I(z) = P_I$ and there would be no value of z for which $P_I^2(z) = 0$ and thus no associated singularity. The fact that Eq. (4.12) may only contain a subset of factorization channels should not trouble you in the least—irrespective of the choice of momentum shift, Cauchy's theorem guarantees that the final result will be the correct amplitude. On the contrary, by shifting as few legs as possible we obtain a recursion relation involving fewer factorization channels, thus resulting in a more compact expression for the amplitude.

4.3   Boundary Terms

The recursion relation in Eq. (4.12) recasts *any* tree-level on-shell amplitude in terms of lower-point on-shell amplitudes and a boundary term, B_∞, whose value hinges crucially on the form of the momentum shift. Much like human beings, momentum shifts are organized into two distinct and easily discernible types: good ($B_\infty = 0$) and bad ($B_\infty \neq 0$). For a good momentum shift, the corresponding amplitude is constructible purely from lower-point on-shell data. Furthermore, we say that a theory is on-shell constructible if there exists a momentum shift for which $B_\infty = 0$ for every tree-level scattering amplitude[4]. As we will soon see, this is criterion is satisfied in an enormous class of theories.

There exist several methods for evaluating the boundary term. The first [44] is to simply use a known representation of the amplitude, *e.g.* the MHV expansion [48]. A second approach involves a return to the action. Here the key insight is to realize that the large z limit coincides with a particular high energy limit where the shifted legs behave like hard particles scattering in a background of soft quanta [49, 50]. To calculate B_∞ one then computes the amplitude for hard particle scattering in the background field method. A third way—more in line with the underlying philosophy of the amplitudes program—is to derive B_∞ purely from general principles like dimensional analysis, Lorentz invariance, and locality [51].

Before continuing, let us comment briefly on *why* it is that certain amplitudes are not on-shell constructible. The classic example of this is the four-particle amplitude of ϕ^4 theory, $A(1234) = \lambda$. Because this amplitude is momentum independent it will not vanish at large z for *any* momentum shift. Naively, this sounds like a spectacular failure: on-shell methods cannot even build the simplest imaginable amplitude!

However, such a criticism is tremendously unfair because it holds on-shell recursion relations to a standard not held even by Feynman diagrams. In particular, $A(1234)$ arises from a Feynman vertex which is simply *assumed* in the action formalism. That is, the ϕ^4 vertex is primordial and unrelated by symmetry to lower order operators. Hence, $A(1234)$ is not related to the three-particle amplitude by factorization and should be included as a *seed* for the on-shell recursion. On-shell recursion can then be used to construct all amplitudes with greater than four particles in ϕ^4 theory. The general lesson

---

[4] In some circumstance one can relate B_∞ to other physical data [47], though we will not consider this possibility here.

here is that interaction vertices which are unrelated by symmetry to lower-order processes are not on-shell constructible and this should be completely unsurprising.

4.4 BCFW Recursion

The first incarnation of on-shell recursion, pioneered by BCFW [43, 44], remains to this day unparalleled in its simplicity. BCFW defined a shift of two external legs given by Eq. (4.1) where

$$q_1 = -q_2 = \lambda_1 \tilde{\lambda}_2. \tag{4.13}$$

In terms of spinor helicity variables, this shift is elegantly written as

$$\tilde{\lambda}_1 \rightarrow \tilde{\lambda}_1 + z\tilde{\lambda}_2 \tag{4.14}$$
$$\lambda_2 \rightarrow \lambda_2 - z\lambda_1, \tag{4.15}$$

which automatically preserves Eq. (4.3) and Eq. (4.4). Since the momentum shift only involves particles 1 and 2, the sum over I in Eq. (4.12) only involves factorization channels in which particles 1 and 2 appear on opposite sides of the channel.

Remarkably, the boundary term B_∞ vanishes both in gauge theory and gravity as long as the momentum shift is applied to external legs with the appropriate configuration of helicities [44, 49]. Presented here without proof, the boundary term is

$$A(1^-2^- \cdots), A(1^-2^+ \cdots), A(1^+2^+ \cdots) \quad \Rightarrow \quad B_\infty = 0 \tag{4.16}$$
$$A(1^+2^- \cdots) \quad \Rightarrow \quad B_\infty \neq 0, \tag{4.17}$$

depending on the helicity configuration of particles 1 and 2. At first glance it may seem peculiar that B_∞ is not invariant under the exchange of plus and minus helicities. However, this is purely an artifact of Eq. (4.15), which picks out a chirality between particles 1 and 2.

Fortunately, in the case of YM theory it is always possible to choose a set of adjacent legs whose helicities match one of the good momentum shifts in Eq. (4.17). This is true because every nontrivial tree-level amplitude in YM theory has at least one plus leg and one minus leg. We thus conclude that YM theory is on-shell constructible via BCFW recursion.

A corollary of this result is that the tree-level S-matrix of YM theory can be built purely from the three-particle amplitudes we derived in Eq. (3.10). This should come as a shock—after all, the YM quartic term is critically necessary in order to maintain gauge invariance of the action. However, on-shell recursion plainly demonstrates that this term carries *zero* physical content beyond the cubic interaction vertex.

Remarkably, the BCFW recursion relations are also applicable to graviton scattering amplitudes [52]. In this case the boundary term is

$$A(1^{--}2^{--}\cdots), A(1^{--}2^{++}\cdots), A(1^{++}2^{++}\cdots) \quad \Rightarrow \quad B_\infty = 0 \qquad (4.18)$$
$$A(1^{++}2^{--}\cdots) \quad \Rightarrow \quad B_\infty \neq 0. \qquad (4.19)$$

The fact that graviton amplitudes vanish at large z for any momentum shift is actually quite remarkable given the infamously poor ultraviolet behavior of gravity. Despite these intuitions, gravity is closely analogous to gauge theory and similarly on-shell constructible.

Just as in YM theory, the tree-level graviton S-matrix can be reduced via on-shell recursion to the primordial three-particle amplitudes we derived earlier in Eq. (3.13). Here we learn that the *infinite tower* of higher order graviton vertices literally serve no purpose but to manifest diffeomorphism invariance at the level of the action.

4.5   All-Line Recursion

We could have foreseen that YM theory and gravity are on-shell constructible. After all, both theories have highly constrained interactions controlled by a single coupling constant. What *is* surprising is that on-shell constructibility is a property of *any renormalizable QFT in four dimensions*. Conveniently, this tremendously large class of theories also happens to have real world relevance.

It is possible to derive on-shell recursion relations for a generic QFT using an all-line holomorphic momentum shift [53, 51],

$$q_i = c_i \eta \tilde{\lambda}_i, \qquad (4.20)$$

where η is an arbitrary reference spinor and c_i are constants with each external leg. This is a generalization of the three-line momentum shift devised by Risager [54] to prove the MHV vertex expansion [48]. Like the reference

spinors used to define external polarization vectors, η will appear in intermediate steps in the recursion but will cancel in any physical answers. By writing Eq. (4.20) as a shift of the holomorphic spinors,

$$\lambda_i \to \lambda_i + z c_i \eta, \tag{4.21}$$

we see that Eq. (4.4) is automatically satisfied while Eq. (4.3) implies

$$\sum_i^n c_i \tilde{\lambda}_i = 0. \tag{4.22}$$

Meanwhile, an anti-holomorphic momentum shift is defined in the obvious way by swapping the holomorphic and anti-holomorphic spinors in Eq. (4.20), Eq. (4.21), and Eq. (4.22).

With nothing more than little group covariance and dimensional analysis [51], one can show that the large z behavior of an all-line shift is

$$\lim_{z \to \infty} A_n(z) = z^m, \tag{4.23}$$

where the exponent of the fall off is

$$m = \begin{cases} \frac{1}{2}(4 - n - [g] + h), & \text{holomorphic shift} \\ \frac{1}{2}(4 - n - [g] - h), & \text{anti-holomorphic shift} \end{cases} \tag{4.24}$$

Here h is the total helicity of the external particles and $[g]$ denotes the mass dimension of all the coupling constants in the amplitude. For the case of a renormalizable theory, $[g] \geq 0$, so

$$m \leq \begin{cases} \frac{1}{2}(4 - n + h), & \text{holomorphic shift} \\ \frac{1}{2}(4 - n - h), & \text{anti-holomorphic shift} \end{cases} \tag{4.25}$$

For $n > 4$, the boundary term in Eq. (4.12) vanishes for holomorphic shifts when $h \leq 0$ and for anti-holomorphic shifts when $h \geq 0$. Since this accounts for every possibility, all amplitudes with $n > 4$ particles are on-shell constructible in a renormalizable QFT. Amplitudes with $n \leq 4$ are the seed amplitudes for this recursion.

In practice, all-line recursion is substantially more complicated to implement than BCFW. Since all the momenta are shifted, the corresponding recursion relation will involve contributions from *every* factorization channel

of the amplitude. Furthermore, the resulting expressions carry spurious dependence on the reference spinor η, thus obscuring the underlying Lorentz invariance of the final answer.

On the other hand, all-line recursion offers the conceptual victory of elevating on-shell methods to the same footing as the tree-level action for all renormalizable QFTs. Admittedly, these results are confined to four dimensions but recall that so too are many of your close friends and family. That is to say, these methods provide a purely on-shell definition of our actual laws of physics, *i.e.* the standard model plus gravity.

5   Infrared Structure

The previous sections detail how physical properties like dimensional analysis, Lorentz invariance, and locality dictate the tree-level S-matrices of an enormous class of theories. Clearly, our methods have the most horsepower in theories either involving renormalizable interactions or particles with spin. Notably missing from this list are *nonrenormalizable* theories of scalars, including numerous effective field theories (EFTs) which are critically important for low-energy phenomena. The archetypal example of an EFT is the nonlinear sigma model (NLSM) [55, 56, 57], which famously describes the universal dynamics of spontaneous symmetry breaking. This theory plays an important role in the strong interactions [58] and electroweak physics [59, 60].

Unfortunately, there are significant obstacles to extending on-shell methods to scalar EFTs. First of all, scalars are singlets of the little group, so the associated considerations are simply not useful for constraining the amplitude. Moreover, the inherent nonrenormalizability of EFTs exacerbates the issue of boundary terms in on-shell recursion. Of course, this second concern is not completely warranted since the same is true of gravity, where we have seen that poor high energy behavior is unimportant.

A greater impediment is that scalar EFTs simply require *more* physical input than just dimensional analysis, Lorentz invariance, and locality. In the textbook formulation of EFTs, the additional ingredient is nonlinearly realized symmetry, *e.g.* chiral symmetry in the NLSM. Unlike gauge and diffeomorphism invariance, this is *genuine* global symmetry with physical consequences for the S-matrix.

Unfortunately, the concepts of nonlinearly realized symmetry, coset spaces, and related paraphernalia are inextricably tied to the notion of an action.

Unmoored from these cherished principles, how can we possible construct the S-matrix of an EFT? As we will see, the path forward is to introduce additional physical information which is admissible in the language of on-shell scattering amplitudes: the soft limit.

The soft limit of an amplitude is defined by sending the momentum of an external leg to zero. To characterize the behavior in the soft limit, we define a soft degree σ [61, 62],

$$\lim_{p \to 0} A(p) \propto p^\sigma, \tag{5.1}$$

which is integer-valued for any tree-level amplitude. When the right-hand side of this equation involves a structure which is universal, *e.g.* zero or a known factor times a lower-point amplitude, it is said that the amplitude satisfies a soft theorem.

There is a long prehistory of deriving soft theorems from an action using chiral symmetry or gauge and diffeomorphism invariance. From this perspective the soft theorems are a *byproduct* of the action. In the subsequent sections, we implement an amusing inversion of this logic: by *assuming* the soft theorems we can derive the S-matrix of the EFT and hence all of its properties. This is a notable reversal of the usual reductionist approach to high energy physics. Eschewing a top-down perspective grounded in symmetry and unification in the ultraviolet, we pursue a bottom-up definition of these theories gleaned from the deep infrared.

5.1   Weinberg Soft Theorems

The prototypical soft theorem was discovered in the seminal work of Weinberg [63, 64] on gauge theory and gravity. As we have shown, since the these theories are on-shell constructible by dimensional analysis, Lorentz invariance, and factorization, these soft theorems cannot encode any additional information beyond this. Nevertheless, the case of gauge theory and gravity will serve as an instructive warmup.

Gauge theory and gravity amplitudes are actually *singular* in the soft limit because soft particles can interact with hard particles through a primordial cubic vertex. Since the soft particle imparts negligible momentum, the hard particle propagator adjacent to the interaction is nearly on-shell, yielding a singularity. Hence, the leading contribution to the soft limit of gauge theory and gravity enters at $\sigma = -1$.

First, let us consider first the case of quantum electrodynamics (QED). As proven by Weinberg [64], the soft limit of a photon of momentum p_μ and polarization e_μ is

$$\lim_{p \to 0} A_{n+1} = e_\mu S^\mu A_n, \tag{5.2}$$

where S^μ is a universal soft function that is *independent* of the hard process. On general grounds we know that every term in S^μ should have a propagator pole together with one power of momentum from the cubic gauge interaction vertex. Indeed, the Weinberg soft factor is

$$S^\mu = \sum_i^n q_i \frac{p_i^\mu}{p_i p}, \tag{5.3}$$

where q_i is the coupling constant associated with the cubic vertex attaching to each hard particle. By the Ward identity, the soft theorem should be invariant under $e_\mu \to e_\mu + \alpha p_\mu$, so

$$\left(\sum_i^n q_i \right) A_n = 0. \tag{5.4}$$

Here q_i should be interpreted as the electric charge of each hard particle, whose sum is conserved by the soft theorem[63].

An analogous story holds for gravitons,

$$\lim_{p \to 0} A_{n+1} = e_{\mu\nu} S^{\mu\nu} A_n. \tag{5.5}$$

In this case, every term in $S^{\mu\nu}$ has a propagator pole together with two powers of momenta from the cubic vertex with the graviton, yielding

$$S^{\mu\nu} = \sum_i^n \kappa_i \frac{p_i^\mu p_i^\nu}{p_i p}. \tag{5.6}$$

The soft theorem should be invariant under a diffeomorphism transformation, $e_{\mu\nu} \to e_{\mu\nu} + \alpha_\mu p_\nu + \alpha_\nu p_\mu$, so we obtain

$$\left(\sum_i^n \kappa_i p_i^\mu \right) A_n = 0. \tag{5.7}$$

We are already familiar with an equation of this form—total momentum conservation implies that $\sum_i^n p_i^\mu = 0$. Any additional constraint on the external momenta would then require that the *hard* particles in the amplitude conform to some special kinematic configuration. As this is plainly impossible, we deduce that

$$\kappa_i = \kappa, \tag{5.8}$$

so the coupling strength of gravity is *universal* [63]. Thus, we have arrived at the celebrated equivalence principle of Einstein, derived here without the ethical pitfalls of hurling unsuspecting test subjects into the dead of space in an elevator.

As one might anticipate, there exist soft theorems for gauge theory and gravity beyond leading order in the soft momentum. In particular, terms which are finite in the soft limit also exhibit various universal structures, *e.g.* for the subleading soft photon theorem, [65, 66] and the subleading [67, 68, 69, 70] and subsubleading soft graviton theorems [70].

5.2 Adler Zero and Beyond

Let us proceed to the subject of amplitudes which *vanish* in the soft limit. A familiar example of this is the NGB, described by the general Lagrangian,

$$\mathcal{L}_{\text{NGB}} = \frac{1}{2}(\partial\phi)^2 + \frac{\lambda_4}{4!}(\partial\phi)^4 + \frac{\lambda_6}{6!}(\partial\phi)^6 + \dots \tag{5.9}$$

Here the derivatively coupled structure of interactions is enforced by the nonlinearly realized shift symmetry of the NGB. Since the field appears everywhere with a derivative, all off-shell Feynman vertices are multilinear in each of the incoming momenta. Consequently, the associated soft degree, $\sigma = 1$, is trivially manifest at the level of the action. Conversely, by *assuming* a soft degree of $\sigma = 1$, we learn very little. This constraint still permits an infinite space of theories parameterized by the coupling constants λ_n.

We can, however, up the ante to $\sigma = 2$. This implies that the amplitude vanishes at $\mathcal{O}(p)$ in the soft limit, which is only possible if there are nontrivial *cancellations* among contributions. Concretely, for each n-particle amplitude, vanishing $\mathcal{O}(p)$ behavior requires that the coupling λ_n of the n-particle vertex is specially related to the couplings at lower order. Iterating this procedure [61], one finds that all the λ_n are determined in terms of the coupling constant

of the leading interaction, λ_4. The resulting coefficients telescope into the Taylor expansion of an *even simpler* Lagrangian,

$$\mathcal{L}_{\text{DBI}} = \frac{1}{2}(\partial\phi)^2 + \frac{\lambda_4}{24}(\partial\phi)^4 + \frac{\lambda_4^2}{144}(\partial\phi)^6 + \frac{5\lambda_4^3}{3456}(\partial\phi)^8 + \dots$$

$$= -\frac{3}{\lambda_4}\sqrt{1 - \frac{\lambda_4}{3}(\partial\phi)^2} + \text{constant}, \tag{5.10}$$

which describes the scalar mode of Dirac-Born-Infeld (DBI) theory. In this theory the scalar labels the coordinate of a brane residing in an extra dimension. The scalar shift symmetry encodes translation symmetry in this direction. Higher-dimensional rotations and boosts are manifested by an additional nonlinearly realized symmetry of the scalar which fixes the peculiar square root form of its interactions. For these reasons we say that DBI theory has an "enhanced soft limit": a soft degree which is higher than naively expected given the number of derivatives per field. Achieving this special property requires *destructive interference* among contributions to the amplitude, which in turn dictates a rigid structure for interactions.

Why stop here? It seems natural to further impose $\sigma = 3$. However, DBI is already defined by a single coupling constant, so setting $\mathcal{O}(p^2)$ soft contributions to zero will simply fix this parameter to zero, yielding a free theory. So there is no interacting $\sigma = 3$ theory with the power counting of one derivative per field. DBI is as simple as it gets.

The concept of an enhanced soft limit is a generalization of the Adler zero [71], which is the observation that soft pions in the NLSM exhibit $\sigma = 1$. At face value this is surprising because the NLSM only has two derivatives per vertex[5]. However, the NLSM encodes nonlinearly realized chiral symmetries which enforce these amplitude zeros in the soft limit. Conversely, Susskind and Frye showed [72] explicitly how the Adler zero can be reverse engineered in order to *derive* the structure of the NLSM.

Through a systematic generalization of the above procedure, one can carve out a theory space of possible Lorentz invariant scalar EFTs [62]. The degree of symmetry in a theory correlates strongly with its soft behavior. For the n-particle amplitude, the soft limit is enhanced when

$$\text{derivatives per particle} = \frac{m}{n} < \sigma = \text{soft degree}, \tag{5.11}$$

---

[5] Earlier we proved that a single scalar with two-derivative interactions has a trivial S-matrix. The NLSM evades this argument by introducing multiple flavors.

where m is the number of derivatives in the n-particle vertex. Imposing this criterion, we derive a special class of "exceptional" theories which exhibit the maximal enhanced soft limit permitted for a self-consistent and nontrivial S-matrix. These exceptional theories include the NLSM ($\sigma = 1$), DBI scalar theory ($\sigma = 2$), and the special Galileon ($\sigma = 3$). The latter was first conjectured [61] and then later realized as an enhanced symmetry point [73] of the original Galileon [74] and as well as an output of the scattering equations [75]. Perhaps unsurprisingly, these scalar EFTs are defined by a *single* coupling constant, consistent with the fact that they cannot be constrained further without becoming trivial. In this sense, these theories are the scalar EFT analogs of YM theory and gravity!

While the soft limit fixes the *form* of interactions in a scalar EFT, what controls the *precise identity* of the nonlinearly realized symmetry? For example, the pions of the NLSM are valued in G/H corresponding to the breaking of a group G down to a subgroup H. Remarkably, this information is still accessible from the S-matrix through the commutator of soft limits [71, 76, 77].

5.3 Soft Recursion

We can capitalize on the soft theorems to derive on-shell recursion relations for scalar EFTs. Here the key insight is to relate the bothersome boundary term of the recursion to the infrared limit of the amplitude.

First, we need a momentum shift which can probe the soft limit of one or more external legs. Here we will focus on the case of a d-dimensional shift [78], though there exist certain versions specialized to four dimensions [46, 79]. In particular, let us consider the soft momentum shift,

$$p_i \rightarrow p_i(1 - za_i), \tag{5.12}$$

where z is the usual deformation parameter and a_i are constants. Since this shift is a rescaling of the momenta, it trivially preserves the on-shell conditions. However, total momentum conservation is only guaranteed if

$$\sum_i^n a_i p_i = 0. \tag{5.13}$$

Unfortunately, this condition cannot always be satisfied in general. If, for example, the total number of external particles is $n = d$, then the momentum vectors in Eq. (5.13) are linearly independent and cannot sum to zero.

For $n = d + 1$ legs the only solution to Eq. (5.13) is $a_i = 1$, corresponding to a trivial rescaling of all the momentum. Because the amplitude is homogenous in the momentum it also rescales, so this shift cannot probe any interesting kinematic regime of the amplitude. Thus, we conclude that a useful momentum shift requires that the number of particles exceed

$$n > d + 1. \tag{5.14}$$

When this condition is satisfied, the a_i can be made distinct so the shift probes the soft limit of leg i as z approaches $1/a_i$. If the soft degree of the theory is σ, then

$$\lim_{z \to 1/a_i} A(z) \propto (z - 1/a_i)^\sigma. \tag{5.15}$$

In other words, the amplitude has multiple zeroes of degree σ.

In order to construct an on-shell recursion relation it will be convenient to define a function $F(z)$ which vanishes on soft kinematics,

$$F(z) = \prod_i^n (1 - a_i z)^\sigma, \tag{5.16}$$

where $F(0) = 1$. Repeating our procedure from before, we simply apply Cauchy's theorem to the deformed amplitude with the slight twist of first dividing by $F(z)$, so

$$A(z = 0) = \frac{1}{2\pi i} \oint_{z=0} dz \, \frac{A(z)}{zF(z)} = -\sum_I \operatorname*{Res}_{z=z_I} \left[\frac{A(z)}{zF(z)} \right] + B_\infty. \tag{5.17}$$

The entire purpose of $F(z)$ is to improve the large z behavior of the amplitude. Naively, this comes at a cost—*new* denominator poles which enter at kinematic configurations which are not factorization channels and whose corresponding residues are completely unrelated to lower-point on-shell amplitudes. However, all of these poles are *by construction* cancelled by the soft zeroes in $A(z)$, so no additional residues need be summed. The only residues which contribute to the right-hand side of the recursion relation are the usual ones related to lower-point objects.

Last but not least is the question of the boundary term, B_∞. For a general nonrenormalizable theory, the high energy behavior is poor and $A(z)$

will blow up at large z. The boundary term vanishes provided $F(z)$ grows even faster, so

$$\lim_{z \to \infty} \frac{A(z)}{F(z)} = 0 \qquad \Rightarrow \qquad B_\infty = 0. \qquad (5.18)$$

Since $F(z)$ is polynomial of degree $n\sigma$, it scales as $z^{n\sigma}$ at large z. On the other hand, $A(z)$ scales at most as z^m where m is the total number of derivatives in the n-particle vertex. So the boundary term vanishes when $\sigma > m/n$, coinciding precisely with the definition of the enhanced soft limit in Eq. (5.11). In conclusion, we discover that all of the exceptional EFTs presented earlier—the NLSM, DBI theory, and the special Galileon—are constructible via on-shell recursion.

6   Color-Kinematics Duality

Up to now our efforts have been centered on reproducing results from QFT purely from on-shell methods. While this is a noble aim in and of itself, the fact remains that nothing we have computed is formally beyond the capabilities of Feynman diagrams plus infinite time or RAM. This is, however, besides the point—the true prize of the modern amplitudes program is the litany of *newly discovered structures*, long hidden in plain sight in theories more than a half a century old.

For the special case of $\mathcal{N} = 4$ SYM, the fruits of this research program has been extraordinary. On-shell tools have excavated a landscape of structures: hidden Yangian and dual conformal symmetries [80, 81, 82, 83, 84], reformulations of the S-matrix in twistor space [85, 86, 22, 87, 88, 89, 90, 91] and Grassmannian space [92, 93, 94], as well as connections to abstract volumes [12, 95] like the amplituhedron [96].

Perhaps less appreciated outside the amplitudes community is that much of this progress also extends far beyond the case of $\mathcal{N} = 4$ SYM. In particular, color-kinematics duality [97] and the scattering equations [98, 99, 75] apply to an enormous class of theories, *e.g.* including gravity and YM theory, but also EFTs like the NLSM! While these structures are simple, provably true facts about the S-matrix, they are no less than sorcery from the perspective of the action.

A comprehensive survey of these new results would fill an entire course. However, given the limitations of time, we focus here on the specific subject

of color-kinematics duality, or Bern-Carrasco-Johansson (BCJ) duality [97]. Epitomized by the word-algebraic equation, "graviton = gluon^2", BCJ is a miraculous duality that interchanges the color and kinematic structures of the S-matrix. Fundamentally, BCJ duality has two basic elements:

- *Kinematic Algebra.* Amplitudes can be rearranged so that their kinematic structures satisfy a kinematic analog of the Jacobi relations. In this basis, color and kinematic structures are dual in the sense that they satisfy the exact same algebraic relations.

- *Double Copy.* Amplitudes in the dual form can be systematically "squared" to generate new amplitudes in *other* theories. This construction generates a web of relations of the form "graviton = gluon^2" and "Galileon = pion^2".

BCJ duality is proven at tree-level [100] and vigorously checked at higher loop [101, 102, 103, 104, 105, 106, 107]. Remarkably, it is also a property of numerous QFTs.

6.1 Color Structure

Before moving forward it will be necessary to review the fundamentals of color in YM theory[6]. For additional details, we refer the reader to several excellent reviews [108, 26, 109] where this topic is covered in more depth.

For concreteness, let us consider YM theory with an $SU(N)$ gauge group. The generators of the Lie algebra satisfy

$$\text{Tr}(T_a T_b) = \delta_{ab} \qquad \text{and} \qquad [T_a, T_b] = i f_{abc} T_c, \tag{6.1}$$

where the trace runs over the fundamental representation. Here we have defined the structure constant,

$$i f_{abc} = \text{Tr}([T_a, T_b] T_c), \tag{6.2}$$

which is cyclically invariant and antisymmetric in all of its indices. We will make frequent use of the completeness relation for $SU(N)$,

$$(T_a)_{ij}(T_a)_{kl} = \delta_{ik}\delta_{jl} - \frac{1}{N}\delta_{ij}\delta_{kl}. \tag{6.3}$$

---

[6] Note that our discussion will also apply to *flavor* structures in the NLSM.

To mitigate the proliferation of $\sqrt{2}$ factors in amplitudes, our normalization convention differs slightly from the one typically used in textbooks.

A gluon scattering amplitude is composed of kinematic factors multiplying color structures built from products of the structure constant, f_{abc}. At tree-level, it is straightforward to reduce all such color structures into *single trace* objects. As a concrete example, consider the Feynman diagram for s-channel gluon exchange in four-particle scattering. The associated color factor is

$$f_{abe}f_{cde} = -\text{Tr}([T_a, T_b]T_e)\text{Tr}([T_c, T_d]T_e) \tag{6.4}$$
$$= -\text{Tr}([T_a, T_b][T_c, T_d]) + \mathcal{O}(1/N). \tag{6.5}$$

By applying the completeness relation in Eq. (6.3), we can trivially rewrite this color structure in terms of single trace terms like $\text{Tr}(T_a T_b T_c T_d)$, plus $1/N$ corrections. Crucially, these $1/N$ corrections actually vanish for tree-level gluon amplitudes. To see why, realize that the $1/N$ terms serve only subtract the trace components of the generators. Adding back in these traces yields the completeness relation for $U(N)$,

$$(T_a)_{ij}(T_a)_{kl} \overset{U(1)}{=} \delta_{ik}\delta_{jl}, \tag{6.6}$$

where all $1/N$ contributions disappear. Physically, the extension from $SU(N)$ to $U(N)$ merely introduces a multiplet of $U(1)$ photons to the theory. However, at the renormalizable level, photons decouple from all gluon interactions due to gauge invariance. So all tree-level amplitudes involving the photon will vanish and we can consistently drop the $1/N$ factors.

Every tree-level amplitude in YM theory has a "color decomposition" into a sum over single trace color structures times kinematic functions,

$$A(1_{a_1}^{h_1} \cdots n_{a_n}^{h_n}) = \sum_{\sigma \in S_n/\mathbb{Z}_n} \text{Tr}(T_{\sigma(a_1)} \cdots T_{\sigma(a_n)}) A(\sigma(1^{h_1}) \cdots \sigma(n^{h_n})). \tag{6.7}$$

The kinematic functions multiplying each color trace are gauge invariant, physical objects known as the "color-ordered" or "partial" amplitudes. In a criminal abuse of notation, we will use the same symbol for both the full and color-ordered amplitude, distinguished only by the presence of color subscripts in the particle labels of the former.

Color-ordered amplitudes are simpler than the full amplitudes. The reason is that these objects have by definition an *intrinsic ordering* of the exter-

nal legs, so their singularity structure is restricted. In particular, factorization channels can only arise from *adjacent* sets of momenta, so for instance $1/s_{123}$ can appear but not $1/s_{135}$.

As an example, take the color decomposition of the four-particle gluon scattering amplitude,

$$
\begin{align}
A(1_a^- 2_b^- 3_c^+ 4_d^+) = \quad & \mathrm{Tr}(T_a T_b T_c T_d)\, A(1^- 2^- 3^+ 4^+) \tag{6.8} \\
+ & \mathrm{Tr}(T_b T_c T_a T_d)\, A(2^- 3^+ 1^- 4^+) \tag{6.9} \\
+ & \mathrm{Tr}(T_c T_a T_b T_d)\, A(3^+ 1^- 2^- 4^+) + \dots \tag{6.10}
\end{align}
$$

Here the color-ordered amplitude is

$$
A(1^- 2^- 3^+ 4^+) = \frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle}. \tag{6.11}
$$

Since the color-ordered amplitude presumes a fixed ordering of the external legs, all of its singularities are in adjacent momenta. The same is true of the Park-Taylor formula in Eq. (2.10).

The color decomposition in Eq. (6.7) is a function of $(n-1)!$ trace structures. However, there are only $(n-2)!$ *independent* color structures spanned by a basis of the form

$$
\left[\tau_{\sigma(a_2)} \cdot \tau_{\sigma(a_3)} \cdots \tau_{\sigma(a_{n-2})} \cdot \tau_{\sigma(a_{n-1})} \right]_{a_1 a_n} \tag{6.12}
$$

where legs 1 and n have been fixed and $\sigma \in S_{n-2}$ is any permutation of the inner $n-2$ legs. Here the products represent matrix multiplication in the adjoint representation, so $i f_{abc} = [\tau_b]_{ac}$. Any possible string of structure constants can be expressed as a linear combination of the above objects via the Jacobi identities. This basis is sometimes called the "half-ladder" basis, in reference to the structure of color index contractions when expressed in terms of f_{abc} factors.

Using the half-ladder basis one can construct an *even more minimal* color decomposition of the YM amplitude,

$$
A(1_{a_1}^{h_1} \cdots n_{a_n}^{h_n}) = \sum_{\sigma \in S_{n-2}} \left[\tau_{\sigma(a_2)} \cdots \tau_{\sigma(a_{n-1})} \right]_{a_1 a_n} A(1^{h_1} \sigma(2^{h_2}) \cdots \sigma(n^{h_{n-1}}) n^{h_n}).
$$

$$\tag{6.13}$$

known as the DDM basis [110]. Note that the number of independent color-ordered amplitudes has decreased from $(n-1)!$ to $(n-2)!$, consistent with

the Kleiss-Kuijf relations [111]. As an example, the DDM decomposition of the four-particle YM amplitude is

$$A(1_a^- 2_b^- 3_c^+ 4_d^+) = [\tau_b \cdot \tau_c]_{ad} \, A(1^- 2^- 3^+ 4^+) + [\tau_c \cdot \tau_b]_{ad} \, A(1^- 3^+ 2^- 4^+), \quad (6.14)$$

which is even more compact than Eq. (6.10).

6.2   A Simple Example

Armed with a basic understanding of color management, we are finally equipped to discuss color-kinematics duality. To be concrete, consider the simplest example of the four-particle amplitude in YM theory. As we saw earlier, this object appears with three possible color structures,

$$c_s = f_{abe} f_{cde} \qquad (6.15)$$
$$c_t = f_{bce} f_{ade} \qquad (6.16)$$
$$c_u = f_{cae} f_{bde}, \qquad (6.17)$$

corresponding to the s-, t-, and u-channel exchange diagrams. Previously, we showed how factorization implies the Jacobi identity, $c_s + c_t + c_u = 0$. Note the similarity of this equation to the *kinematic* relation, $s + t + u = 0$. As we will soon see, this is not mere coincidence!

Starting from *any* representation of the four-particle scattering amplitude, one can always rearrange terms into to the form

$$A(1_a 2_b 3_c 4_d) = \frac{c_s n_s}{s} + \frac{c_t n_t}{t} + \frac{c_u n_u}{u}. \qquad (6.18)$$

While the Feynman diagram representation of the amplitude obviously has cubic exchange diagrams in this form, the same is not true for the quartic vertex contributions. However, these contact terms can be massaged into a cubic form simply by canceling denominator poles with additional numerator factors. That is, any quartic vertex can be recast as a cubic exchange diagram by inserting the equation,

$$1 = \frac{s}{s} = \frac{t}{t} = \frac{u}{u}, \qquad (6.19)$$

whose proof we leave as an exercise for the reader.

Of course, any such rearrangement of contact terms will not be unique, but this is exactly the point. By exploiting this freedom, BCJ conjectured

that there *always* exists a form of the amplitude in which the numerators satisfy the Jacobi identity,

$$n_s + n_t + n_u = 0. \tag{6.20}$$

Since the BCJ numerators satisfy the same identities as the color factors, it is said that this representation manifests color-kinematics duality. That a BCJ form should exist may seem like a miracle, but we have at our disposal an immense redundancy. That is, for various choices of gauge and field basis, the distribution of terms between the s-, t-, and u-channel topologies will change. Under this broad set of "generalized gauge transformations", the numerators of the amplitude can *only* shift by

$$n_s \to n_s + \alpha s \tag{6.21}$$

$$n_t \to n_t + \alpha t \tag{6.22}$$

$$n_u \to n_u + \alpha u, \tag{6.23}$$

so that the physical amplitude remains invariant,

$$A_4 \to A_4 + \alpha \left(c_s + c_t + c_u \right) = A_4, \tag{6.24}$$

by the color Jacobi identities. Here α is an arbitrary, possibly non-local function of the external kinematics parameterizing the space of generalized gauge transformations.

Because the four-particle amplitude is so simple it is easy to derive explicit expressions for the BCJ numerators. From our earlier discussion we argued that the four-particle amplitudes takes the form

$$A(1_a^- 2_b^- 3_c^+ 4_d^+) = \langle 12 \rangle^2 [34]^2 \left(\frac{c_{st}}{st} + \frac{c_{tu}}{tu} + \frac{c_{us}}{us} \right), \tag{6.25}$$

where c_{st}, c_{tu}, and c_{us} are dimensionless constants. Proper factorization implies that these constants are related to the color structures by

$$c_{st} - c_{us} = c_s \tag{6.26}$$

$$c_{tu} - c_{st} = c_t \tag{6.27}$$

$$c_{us} - c_{tu} = c_u. \tag{6.28}$$

Comparing Eq. (6.18) and Eq. (6.25), we obtain

$$\begin{aligned} A(1_a^- 2_b^- 3_c^+ 4_d^+) &= \frac{(c_{st} - c_{us})n_s}{s} + \frac{(c_{tu} - c_{st})n_t}{t} + \frac{(c_{us} - c_{tu})n_u}{u} \\ &= \frac{c_{st}(tn_s - sn_t)}{st} + \frac{c_{tu}(un_t - tn_u)}{tu} + \frac{c_{us}(sn_u - un_s)}{us}. \end{aligned} \tag{6.29}$$

Matching term by term, we obtain three equations for the BCJ numerators,

$$tn_s - sn_t = \langle 12 \rangle^2 [34]^2 \tag{6.30}$$

$$un_t - tn_u = \langle 12 \rangle^2 [34]^2 \tag{6.31}$$

$$sn_u - un_s = \langle 12 \rangle^2 [34]^2, \tag{6.32}$$

which as a matrix equation becomes

$$\begin{pmatrix} t & -s & 0 \\ 0 & u & -t \\ -u & 0 & s \end{pmatrix} \begin{pmatrix} n_s \\ n_t \\ n_u \end{pmatrix} = \langle 12 \rangle^2 [34]^2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{6.33}$$

The above matrix is not invertible because it has vanishing determinant. This is not surprising because $(n_s, n_t, n_u) = (s, t, u)$ is a zero eigenvector of the matrix, corresponding to generalized gauge invariance. No matter though—the fact that the matrix is not invertible simply implies that there is a family of solutions. Solving for n_s and n_t in terms of n_u, we find

$$n_s = -\frac{\langle 12 \rangle^2 [34]^2}{u} + \frac{n_u s}{u} \tag{6.34}$$

$$n_t = \frac{\langle 12 \rangle^2 [34]^2}{u} + \frac{n_u t}{u}, \tag{6.35}$$

for arbitrary n_u. Remarkably, the kinematic Jacobi identity, $n_s + n_t + n_u = 0$, is satisfied for any n_u as a direct consequence of $s + t + u = 0$. Thus, *any* representation of the four-particle amplitude provides a viable set of BCJ numerators. This is a very special property of four-particle kinematics.

As a double check, one can compute the color-ordered amplitude,

$$A(1^- 2^- 3^+ 4^+) = -\frac{n_s}{s} + \frac{n_t}{t} = \frac{\langle 12 \rangle^2 [34]^2}{u} \left(\frac{1}{s} + \frac{1}{t} \right) \tag{6.36}$$

$$= -\frac{\langle 12 \rangle^2 [34]^2}{st} = \frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle}, \tag{6.37}$$

which correctly matches the Park-Taylor formula. Note the various signs entering through the color factors.

Because the BCJ numerators satisfy the Jacobi identities, we can think of the latter as *defining* relations for the former. In this sense there is little difference, at least algebraically, between the color and kinematic factors in

the amplitude. Thus, it is natural to simply *interchange* the color factors with kinematic factors in various ways. For instance, by substituting in color for kinematics, we obtain

$$A(1234) = \frac{c_s^2}{s} + \frac{c_t^2}{t} + \frac{c_u^2}{u},\tag{6.38}$$

which is the amplitude for a theory of scalars coupled via trilinear interactions. The form of the color structures indicates that each scalar has two independent adjoint color indices. For obvious reasons this theory is known as the bi-adjoint scalar (BS) theory. Alternatively, we can substitute kinematics for color, yielding

$$A(1^{--}2^{--}3^{++}4^{++}) = \frac{n_s^2}{s} + \frac{n_t^2}{t} + \frac{n_u^2}{u}\tag{6.39}$$

$$= \frac{\langle 12\rangle^4[34]^4(s+t) + n_u^2 st(s+t+u)}{stu^2}\tag{6.40}$$

$$= -\frac{\langle 12\rangle^4[34]^4}{stu},\tag{6.41}$$

which is precisely the four-particle graviton amplitude! This construction is referred to as the "double copy" for obvious reasons. However, it is a bit of a misnomer because one need not replace color factors with the kinematic factors appearing in the same amplitude. One can instead mix and match color and kinematic factors from *different* amplitudes, yielding asymmetric copies which turn out to be equally sensible scattering amplitudes.

6.3   Universality of the Double Copy

The extraordinary structures introduced in the previous sections are apparently ubiquitous. Not only do these concepts extend to higher-point tree- and loop-level amplitudes, they also generalize across a huge range of theories. That is, if you can think of a Lorentz invariant theory with a single coupling constant and a name, chances are it lies on one or another side of a double copy relation.

BCJ duality holds for the n-particle tree-level YM amplitude, which can always be recast into the form

$$A_{\mathrm{YM}} = \sum_i \frac{c_i n_i}{d_i},\tag{6.42}$$

where the sum runs over all cubic topologies and the d_i are the associated products of propagator denominators. As before, all quartic and higher vertices are trivially recast as cubic diagrams by multiplying and then dividing by the appropriate propagator factors.

Within the full set of cubic topologies, one can always pick out three which differ only by a single propagator factor. This is a higher-point generalization of the classification of s-, t-, and u-channel diagrams. BCJ duality implies that the numerators can always be chosen so that for *every* triplet of topologies i, j, k the associated color and kinematic factors satisfy

$$c_i + c_j + c_k = 0 \qquad \text{and} \qquad n_i + n_j + n_k = 0. \tag{6.43}$$

In this BCJ form we can substitute kinematics for color to obtain the amplitude of BS theory,

$$A_{\text{BS}} = \sum_i \frac{c_i^2}{d_i}, \tag{6.44}$$

or substitute color for kinematics to obtain the graviton amplitude,

$$A_{\text{G}} = \sum_i \frac{n_i^2}{d_i}. \tag{6.45}$$

Hence, the double copy construction is a rather elegant method for leapfrogging from gauge theory directly to gravity. Moreover, the BCJ construction applies in *general* dimensions, so it does not hinge on any special properties of four-dimensional kinematics.

As noted, BCJ duality extends beyond the case of gauge theory and gravity. For instance, one can express the amplitudes of the NLSM as

$$A_{\text{NLSM}} = \sum_i \frac{c_i r_i}{d_i}, \tag{6.46}$$

where r_i are kinematic numerators satisfying $r_i + r_j + r_k = 0$, and the c_i are the same objects defined earlier except interpreted physically as *flavor* structures. Again substituting color for kinematics, we obtain

$$A_{\text{SG}} = \sum_i \frac{r_i^2}{d_i}, \tag{6.47}$$

	BS	NLSM	YM
BS	BS	NLSM	YM
NLSM	NLSM	SG	BI
YM	YM	BI	G

Fig. 1: Multiplication table of QFTs, including bi-adjoint scalar (BS) theory, the nonlinear sigma model (NLSM), Yang-Mills (YM) theory, the special Galileon (SG), Born-Infeld (BI) theory, and gravity (G).

which is the amplitude for the special Galileon. Taking an asymmetric product of the YM and NLSM numerators, we obtain

$$A_{\mathrm{BI}} = \sum_i \frac{n_i r_i}{d_i}, \tag{6.48}$$

which is the amplitude of BI theory. These relations are concisely summarized in the "multiplication table" of BCJ double copies depicted in Fig. 1.

The universality of the double copy is not yet fully understood. A partial explanation can be found in the scattering equations of Cachazo, He, and Yuan [98, 99, 75], who proposed an extraordinarily compact construction for the S-matrices of *all* the theories in Fig. 1. That this is even possible is a manifestation of a hidden unity among these theories, instantiated by a set of simple relations [112] which extract all their S-matrices from that of gravity.

Last but not least, the double copy also generalizes straightforwardly to the case of supersymmetric theories. For instance, the product of $\mathcal{N} = 4$ SYM with a second copy of $\mathcal{N} = 0, 1, 2$, or 4 SYM is simply $\mathcal{N} = 4, 5, 6$, or 8 supergravity! Other combinations of gauge theories produce non-pure supergravities with additional matter. These relations have been critical for alleviating some of the immense technical challenges in studying the ultraviolet properties of $\mathcal{N} = 8$ supergravity [113, 35].

6.4   Beyond Scattering Amplitudes

Color-kinematics duality is especially noteworthy because it quite possibly extends *beyond* the context of the perturbative S-matrix. In particular, an apparent vestige of the duality has been discovered in *classical* solutions in gauge theory and gravity. In its simplest incarnation, "black hole = charge^2"

[114, 115, 116, 117], the Schwarzschild solution of general relativity is shown to be dual to the Coulomb potential of electrodynamics. Since then, similar double copy structures have also been found in classical radiation [118, 119] and other curved backgrounds [120, 121].

In hindsight, a classical manifestation of the double copy is not unexpected—after all, tree-level amplitudes are *literally* the solutions to classical equations of motion initialized by plane wave perturbations at infinity [122]. Indeed, Duff [123] showed long ago how to construct the Schwarzschild solution by an appropriate convolution of perturbative scattering amplitudes. Nevertheless, the above examples elevate the notion of a nonperturbative double copy from hallucinatory to tantalizingly possible.

6.5 Physical Origins

While no one disputes that color-kinematics duality *is* true, the question remains, *why* it is true? That is, despite myriad mathematical proofs of its veracity, the *physical* principle behind the duality remains elusive. Where in the textbook derivation of QFT are the secret instructions to rearrange diagrams into cubic structures satisfying kinematic Jacobi identities and so on and so forth?

It is tempting to contemplate an explanation for color-kinematics duality purely within the confines of standard QFT—ideally, from an explicit action. In the best of all possible worlds, such an action would manifest *all* the symmetries of the double-copy construction: *i*) a parity exchanging color and kinematic algebras, and *ii*) a pair of independent Lorentz symmetries. The latter is a trivial consequence of the fact that the double copy amplitudes are built from *separately* Lorentz invariant numerators.

Recent years have seen incremental progress towards an action-level understanding of these structures. For instance, Monteiro and O'Connell [124] beautifully demonstrated how the self-dual sector of YM theory squares into self-dual gravity. In this formulation the cubic Feynman vertices are the structure constants of the kinematic algebra. Unfortunately, this has yet to generalized to YM beyond the self-dual limit or outside of four dimensions.

More recently, color-kinematics duality has been manifested explicitly in a cubic reformulation of the NLSM action [125]. The associated Feynman diagrams automatically produce expressions compliant with the BCJ form, while the Jacobi identities arise from an off-shell symmetry closely connected to Lorentz invariance. Squaring the *action* then yields a new action for the

special Galileon which manifests a twofold Lorentz invariance. While it is also possible to rewrite the Einstein-Hilbert action in a way that manifests a twofold Lorentz invariance [28, 126], the connection to color-kinematics duality remains murky.

7 Outlook

Upon the discovery of the muon, Isidor Rabi famously quipped "who ordered that?" Appropriate to the times, those words affirm an implicit expectation that there are many ways that the world *could* have been, but only one way that it *is*. In one sense this is true. After all, the standard model exhibits a baroque pattern of masses and couplings.

On the other hand, the on-shell formulation of QFT demonstrates how dynamical structures like the *form* of interactions in gravity, gauge theory, and effective field theories are fixed *uniquely and directly* by physical principles. In this respect, nature's menu is prix fixe. By relinquishing the action principle—together with its perks and pitfalls—we are finally able to see some of the underlying structures which have lain dormant despite decades of study. There is still much to understand!

Acknowledgements

Further Reading

There are a many of excellent introductory treatments of scattering amplitudes in review articles and textbooks. A very partial list includes:

- *"Gauge and Gravity Amplitudes Relations"* [109], by Carrasco

- *"Calculating Amplitudes Efficiently"* [108], *"Scattering Amplitudes: the Most Perfect Microscopic Structures in the Universe"* [128], and *"A Brief Introduction to Modern Amplitudes Methods* [129], by Dixon

- *"Scattering Amplitudes"* [130] and *"Scattering Amplitudes in Gauge Theory and Gravity"* [26], by Elvang and Huang

- *"Scattering Amplitudes in Gauge Theories"* [131], by Henn and Plefka

- *"Quantum Field Theory and the Standard Model"* [132], by Schwartz

- *"Quantum Field Theory in a Nutshell"* [27], by Zee

Appendix A   Counting Feynman Diagrams

For an n-particle amplitude the corresponding number of Feynman diagrams c_n typically grows quite rapidly with n. In this appendix we derive explicit formulae for c_n in a theory with arbitrary interactions.

Cubic Scalar Theory. To start, consider the case of ϕ^3 theory. We introduce "toy" equations of motion in which the d'Alembertian and the cubic coupling have been set to one and ϕ is simply a function, so

$$\phi = J + \frac{\phi^2}{2!}. \tag{A.1}$$

The perturbative solution to the equations of motion in the presence of a source is in one-to-one correspondence with the generating functional of connected tree amplitudes [122]. Since the toy equations of motion effectively set all propagators and Feynman vertices to one, the corresponding solution simply counts the number of diagrams. So c_n is extracted from the numerical coefficient of the J^{n-1} term in the solution. Solving Eq. (A.4), we obtain

$$\phi(J) = \phi(J) = 1 - \sqrt{1 - 2J} \tag{A.2}$$

$$= J + (1)\frac{J^2}{2!} + (3)\frac{J^3}{3!} + (15)\frac{J^4}{4!} + (105)\frac{J^5}{5!} + \dots, \tag{A.3}$$

where we have chosen the root satisfying $\phi(0) = 0$. We thus correctly conclude that the number of cubic Feynman diagrams is $c_n = (2n - 5)!!$.

Yang-Mills Theory. The same logic applies to theories with higher-degree interaction vertices. For example, YM theory has both cubic and quartic vertices, so the toy equations of motion are

$$A = J + \frac{A^2}{2!} + \frac{A^3}{3!}. \tag{A.4}$$

Like before, we choose the root with $A(0) = 0$, so

$$A(J) = -1 + \frac{3\left(1 - i\sqrt{3}\right)}{2\sqrt[3]{\sqrt{9J^2 + 24J - 11} + 3J + 4}} \tag{A.5}$$

$$+ \frac{1}{2}\left(1 + i\sqrt{3}\right)\sqrt[3]{\sqrt{9J^2 + 24J - 11} + 3J + 4} \tag{A.6}$$

$$= J + (1)\frac{J^2}{2!} + (4)\frac{J^3}{3!} + (25)\frac{J^4}{4!} + (220)\frac{J^5}{5!} + \dots, \tag{A.7}$$

which is in agreement with the counting from known results.

Gravity. Last but not least is the case of gravity, whose toy equations of motion extend to infinite order,

$$h = J + \frac{h^2}{2!} + \frac{h^3}{3!} + \frac{h^4}{4!} + \dots = J + e^h - 1 - h, \tag{A.8}$$

which has a semi-closed form solution,

$$h(J) = \frac{J - 1}{2} - W\left(-e^{\frac{J-1}{2}}/2\right) \tag{A.9}$$

$$= J + (1)\frac{J^2}{2!} + (4)\frac{J^3}{3!} + (26)\frac{J^4}{4!} + (236)\frac{J^5}{5!} + \dots, \tag{A.10}$$

where W is the product logarithm function.

Appendix B Basis of Mandelstam Invariants

In general dimensions, any n-particle amplitude can be expressed as a function of the Mandelstam invariants,

$$\{p_i p_j\}, \qquad \{p_i e_j\}, \qquad \{e_i e_j\}, \tag{B.1}$$

for all $1 \leq i, j \leq n$ where $i \neq j$. While the $\{e_i e_j\}$ are all independent variables, the same cannot be said for $\{p_i p_j\}$ and $\{p_i e_j\}$ because of momentum conservation. On the other hand, one can always prune down to a linearly independent basis. For instance by eliminating p_n via momentum conservation, we can effectively drop all Mandelstam invariants of the form $p_n p_i$ and $p_n e_i$. Furthermore, the on-shell condition for p_n implies

$$0 = \sum_{i}^{n-1} \sum_{j}^{n-1} p_i p_j, \tag{B.2}$$

while the transverse condition on e_n implies

$$0 = \sum_{i}^{n-1} p_i e_n. \tag{B.3}$$

Altogether, these eliminate n of the $n(n-1)/2$ variables in $\{p_i p_j\}$ and n of the $n(n-1)$ variables in $\{p_i e_j\}$. Obviously leaves an immense freedom in choosing which variables to eliminate. However, any full set of linearly independent Mandelstam invariants takes the form

$$\{p_i p_j\}_{\frac{n(n-3)}{2}}, \qquad \{p_i e_j\}_{n(n-2)}, \qquad \{e_i e_j\}_{\frac{n(n-1)}{2}}, \tag{B.4}$$

where the subscripts denote the number of elements in each set.

Depending on the circumstance, different forms of this basis may be appropriate. For example, if one is interested in an n-particle amplitude with adjacent factorization channels, then for odd n, a convenient basis is

$$\{p_i p_j\}_{\frac{n(n-3)}{2}} = \left\{ s_{12}, s_{123}, s_{1234}, \ldots, s_{12\ldots\frac{n-1}{2}} \right\} + \text{cyclic perm}, \tag{B.5}$$

where we have defined the generalized Mandelstam invariant,

$$s_{i_1 i_2 \ldots i_m} = (p_{i_1} + p_{i_2} + \ldots + p_{i_m})^2. \tag{B.6}$$

For the case of even n, a similar basis exists except that the set of n variables of the form $s_{12\ldots\frac{n-1}{2}}$ is replaced with $n/2$ variables of the form $s_{12\ldots\frac{n}{2}}$.

References

[1] V. Bargmann and E. P. Wigner, *Group Theoretical Discussion of Relativistic Wave Equations*, *Proc. Nat. Acad. Sci.* **34** (1948) 211.

[2] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations.* Cambridge University Press, 2005.

[3] S. Mandelstam, *Determination of the pion - nucleon scattering amplitude from dispersion relations and unitarity. General theory,* *Phys. Rev.* **112** (1958) 1344–1360.

[4] R. Haag, *Quantum field theories with composite particles and asymptotic conditions*, *Phys. Rev.* **112** (1958) 669–673.

[5] J. F. Donoghue, E. Golowich and B. R. Holstein, *Dynamics of the standard model, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **2** (1992) 1–540.

[6] D. A. McGady and L. Rodina, *Higher-spin massless S-matrices in four-dimensions*, *Phys. Rev.* **D90** (2014) 084048, [`1311.2938`].

[7] 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) 53–60.

[8] R. Kleiss and W. J. Stirling, *Spinor Techniques for Calculating p anti-p —> W+- / Z0 + Jets*, *Nucl. Phys.* **B262** (1985) 235–262.

[9] Z. Xu, D.-H. Zhang and L. Chang, *Helicity Amplitudes for Multiple Bremsstrahlung in Massless Nonabelian Gauge Theories*, *Nucl. Phys.* **B291** (1987) 392–428.

[10] J. F. Gunion and Z. Kunszt, *Improved Analytic Techniques for Tree Graph Calculations and the G g q anti-q Lepton anti-Lepton Subprocess*, *Phys. Lett.* **161B** (1985) 333.

[11] R. Penrose, *Twistor algebra*, *J. Math. Phys.* **8** (1967) 345.

[12] A. Hodges, *Eliminating spurious poles from gauge-theoretic amplitudes*, *JHEP* **05** (2013) 135, [`0905.1473`].

[13] A. Agarwal, N. Beisert and T. McLoughlin, *Scattering in Mass-Deformed N>=4 Chern-Simons Models*, *JHEP* **06** (2009) 045, [`0812.3367`].

[14] C. Cheung and D. O'Connell, *Amplitudes and Spinor-Helicity in Six Dimensions*, *JHEP* **07** (2009) 075, [`0902.0981`].

[15] S. Caron-Huot and D. O'Connell, *Spinor Helicity and Dual Conformal Symmetry in Ten Dimensions*, *JHEP* **08** (2011) 014, [`1010.5487`].

[16] S. Dittmaier, *Weyl-van der Waerden formalism for helicity amplitudes of massive particles*, *Phys. Rev.* **D59** (1998) 016007, [`hep-ph/9805445`].

[17] C. Schwinn and S. Weinzierl, *SUSY ward identities for multi-gluon helicity amplitudes with massive quarks*, *JHEP* **03** (2006) 030, [`hep-th/0602012`].

[18] R. Boels, *Covariant representation theory of the Poincare algebra and some of its extensions*, *JHEP* **01** (2010) 010, [`0908.0738`].

[19] S. J. Parke and T. R. Taylor, *Gluonic Two Goes to Four*, *Nucl. Phys.* **B269** (1986) 410–420.

[20] S. J. Parke and T. R. Taylor, *An Amplitude for n Gluon Scattering*, *Phys. Rev. Lett.* **56** (1986) 2459.

[21] V. P. Nair, *A Current Algebra for Some Gauge Theory Amplitudes*, *Phys. Lett.* **B214** (1988) 215–218.

[22] E. Witten, *Perturbative gauge theory as a string theory in twistor space*, *Commun. Math. Phys.* **252** (2004) 189–258, [`hep-th/0312171`].

[23] T. He, P. Mitra and A. Strominger, *2D Kac-Moody Symmetry of 4D Yang-Mills Theory*, *JHEP* **10** (2016) 137, [`1503.02663`].

[24] C. Cheung, A. de la Fuente and R. Sundrum, *4D scattering amplitudes and asymptotic symmetries from 2D CFT*, *JHEP* **01** (2017) 112, [`1609.00732`].

[25] P. Benincasa and F. Cachazo, *Consistency Conditions on the S-Matrix of Massless Particles*, `0705.4305`.

[26] H. Elvang and Y.-t. Huang, *Scattering Amplitudes in Gauge Theory and Gravity*. Cambridge University Press, 2015.

[27] A. Zee, *Quantum field theory in a nutshell*. 2003.

[28] C. Cheung and G. N. Remmen, *Twofold Symmetries of the Pure Gravity Action*, *JHEP* **01** (2017) 104, [1612.03927].

[29] G. 't Hooft and M. J. G. Veltman, *One loop divergencies in the theory of gravitation*, *Ann. Inst. H. Poincare Phys. Theor.* **A20** (1974) 69–94.

[30] M. H. Goroff and A. Sagnotti, *The Ultraviolet Behavior of Einstein Gravity*, *Nucl. Phys.* **B266** (1986) 709–736.

[31] M. H. Goroff and A. Sagnotti, *QUANTUM GRAVITY AT TWO LOOPS*, *Phys. Lett.* **160B** (1985) 81–86.

[32] A. E. M. van de Ven, *Two loop quantum gravity*, *Nucl. Phys.* **B378** (1992) 309–366.

[33] Z. Bern, C. Cheung, H.-H. Chi, S. Davies, L. Dixon and J. Nohle, *Evanescent Effects Can Alter Ultraviolet Divergences in Quantum Gravity without Physical Consequences*, *Phys. Rev. Lett.* **115** (2015) 211301, [1507.06118].

[34] Z. Bern, S. Davies, T. Dennen and Y.-t. Huang, *Absence of Three-Loop Four-Point Divergences in N=4 Supergravity*, *Phys. Rev. Lett.* **108** (2012) 201301, [1202.3423].

[35] Z. Bern, S. Davies, T. Dennen and Y.-t. Huang, *Ultraviolet Cancellations in Half-Maximal Supergravity as a Consequence of the Double-Copy Structure*, *Phys. Rev.* **D86** (2012) 105014, [1209.2472].

[36] G. Bossard, P. S. Howe and K. S. Stelle, *Anomalies and divergences in N=4 supergravity*, *Phys. Lett.* **B719** (2013) 424–429, [1212.0841].

[37] G. Bossard, P. S. Howe and K. S. Stelle, *Invariants and divergences in half-maximal supergravity theories*, *JHEP* **07** (2013) 117, [1304.7753].

[38] Z. Bern, S. Davies and T. Dennen, *The Ultraviolet Structure of Half-Maximal Supergravity with Matter Multiplets at Two and Three Loops*, *Phys. Rev.* **D88** (2013) 065007, [1305.4876].

[39] Z. Bern, S. Davies and T. Dennen, *Enhanced ultraviolet cancellations in $\mathcal{N} = 5$ supergravity at four loops*, *Phys. Rev.* **D90** (2014) 105011, [1409.3089].

[40] N. Arkani-Hamed, L. Rodina and J. Trnka, *Locality and Unitarity from Singularities and Gauge Invariance*, 1612.02797.

[41] L. Rodina, *Uniqueness from locality and BCFW shifts*, 1612.03885.

[42] L. Rodina, *Uniqueness from gauge invariance and the Adler zero*, 1612.06342.

[43] R. Britto, F. Cachazo and B. Feng, *New recursion relations for tree amplitudes of gluons*, *Nucl.Phys.* **B715** (2005) 499–522, [hep-th/0412308].

[44] 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, [hep-th/0501052].

[45] F. A. Berends and W. T. Giele, *Recursive Calculations for Processes with n Gluons*, *Nucl. Phys.* **B306** (1988) 759–808.

[46] C. Cheung, C.-H. Shen and J. Trnka, *Simple Recursion Relations for General Field Theories*, *JHEP* **06** (2015) 118, [1502.05057].

[47] B. Feng, J. Wang, Y. Wang and Z. Zhang, *BCFW Recursion Relation with Nonzero Boundary Contribution*, *JHEP* **01** (2010) 019, [0911.0301].

[48] F. Cachazo, P. Svrcek and E. Witten, *MHV vertices and tree amplitudes in gauge theory*, *JHEP* **09** (2004) 006, [hep-th/0403047].

[49] N. Arkani-Hamed and J. Kaplan, *On Tree Amplitudes in Gauge Theory and Gravity*, *JHEP* **04** (2008) 076, [0801.2385].

[50] C. Cheung, *On-Shell Recursion Relations for Generic Theories*, *JHEP* **03** (2010) 098, [0808.0504].

[51] T. Cohen, H. Elvang and M. Kiermaier, *On-shell constructibility of tree amplitudes in general field theories*, *JHEP* **04** (2011) 053, [1010.0257].

[52] P. Benincasa, C. Boucher-Veronneau and F. Cachazo, *Taming Tree Amplitudes In General Relativity*, *JHEP* **11** (2007) 057, [hep-th/0702032].

[53] H. Elvang, D. Z. Freedman and M. Kiermaier, *Proof of the MHV vertex expansion for all tree amplitudes in N=4 SYM theory*, *JHEP* **06** (2009) 068, [0811.3624].

[54] K. Risager, *A Direct proof of the CSW rules*, *JHEP* **12** (2005) 003, [hep-th/0508206].

[55] M. Gell-Mann and M. Levy, *The axial vector current in beta decay*, *Nuovo Cim.* **16** (1960) 705.

[56] S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 1.*, *Phys. Rev.* **177** (1969) 2239–2247.

[57] C. G. Callan, Jr., S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 2.*, *Phys. Rev.* **177** (1969) 2247–2250.

[58] S. Weinberg, *Nonlinear realizations of chiral symmetry*, *Phys. Rev.* **166** (1968) 1568–1577.

[59] T. Appelquist and C. W. Bernard, *Strongly Interacting Higgs Bosons*, *Phys. Rev.* **D22** (1980) 200.

[60] A. C. Longhitano, *Heavy Higgs Bosons in the Weinberg-Salam Model*, *Phys. Rev.* **D22** (1980) 1166.

[61] C. Cheung, K. Kampf, J. Novotny and J. Trnka, *Effective Field Theories from Soft Limits of Scattering Amplitudes*, *Phys. Rev. Lett.* **114** (2015) 221602, [1412.4095].

[62] C. Cheung, K. Kampf, J. Novotny, C.-H. Shen and J. Trnka, *A Periodic Table of Effective Field Theories*, *JHEP* **02** (2017) 020, [1611.03137].

[63] S. Weinberg, *Photons and Gravitons in s Matrix Theory: Derivation of Charge Conservation and Equality of Gravitational and Inertial Mass*, *Phys. Rev.* **135** (1964) B1049–B1056.

[64] S. Weinberg, *Infrared photons and gravitons*, *Phys. Rev.* **140** (1965) B516–B524.

[65] F. E. Low, *Bremsstrahlung of very low-energy quanta in elementary particle collisions*, *Phys. Rev.* **110** (1958) 974–977.

[66] T. H. Burnett and N. M. Kroll, *Extension of the low soft photon theorem*, *Phys. Rev. Lett.* **20** (1968) 86.

[67] D. J. Gross and R. Jackiw, *Low-Energy Theorem for Graviton Scattering*, *Phys. Rev.* **166** (1968) 1287–1292.

[68] R. Jackiw, *Low-Energy Theorems for Massless Bosons: Photons and Gravitons*, *Phys. Rev.* **168** (1968) 1623–1633.

[69] C. D. White, *Factorization Properties of Soft Graviton Amplitudes*, *JHEP* **05** (2011) 060, [`1103.2981`].

[70] F. Cachazo and A. Strominger, *Evidence for a New Soft Graviton Theorem*, `1404.4091`.

[71] S. L. Adler, *Consistency conditions on the strong interactions implied by a partially conserved axial vector current*, *Phys. Rev.* **137** (1965) B1022–B1033.

[72] L. Susskind and G. Frye, *Algebraic aspects of pionic duality diagrams*, *Phys. Rev.* **D1** (1970) 1682–1686.

[73] K. Hinterbichler and A. Joyce, *Hidden symmetry of the Galileon*, *Phys. Rev.* **D92** (2015) 023503, [`1501.07600`].

[74] A. Nicolis, R. Rattazzi and E. Trincherini, *The Galileon as a local modification of gravity*, *Phys. Rev.* **D79** (2009) 064036, [`0811.2197`].

[75] F. Cachazo, S. He and E. Y. Yuan, *Scattering Equations and Matrices: From Einstein To Yang-Mills, DBI and NLSM*, *JHEP* **07** (2015) 149, [`1412.3479`].

[76] S. Weinberg, *Pion scattering lengths*,
Phys. Rev. Lett. **17** (1966) 616–621.

[77] N. Arkani-Hamed, F. Cachazo and J. Kaplan, *What is the Simplest Quantum Field Theory?*, *JHEP* **1009** (2010) 016, [`0808.1446`].

[78] C. Cheung, K. Kampf, J. Novotny, C.-H. Shen and J. Trnka, *On-Shell Recursion Relations for Effective Field Theories*,
Phys. Rev. Lett. **116** (2016) 041601, [`1509.03309`].

[79] F. Cachazo, P. Cha and S. Mizera, *Extensions of Theories from Soft Limits*, *JHEP* **06** (2016) 170, [`1604.03893`].

[80] J. M. Drummond, J. M. Henn and J. Plefka, *Yangian symmetry of scattering amplitudes in N=4 super Yang-Mills theory*,
JHEP **05** (2009) 046, [`0902.2987`].

[81] J. M. Drummond, J. Henn, G. P. Korchemsky and E. Sokatchev, *Dual superconformal symmetry of scattering amplitudes in N=4 super-Yang-Mills theory*, *Nucl. Phys.* **B828** (2010) 317–374,
[`0807.1095`].

[82] N. Berkovits and J. Maldacena, *Fermionic T-Duality, Dual Superconformal Symmetry, and the Amplitude/Wilson Loop Connection*, *JHEP* **09** (2008) 062, [`0807.3196`].

[83] N. Beisert, R. Ricci, A. A. Tseytlin and M. Wolf, *Dual Superconformal Symmetry from AdS(5) x S**5 Superstring Integrability*, *Phys. Rev.* **D78** (2008) 126004, [`0807.3228`].

[84] A. Brandhuber, P. Heslop and G. Travaglini, *A Note on dual superconformal symmetry of the N=4 super Yang-Mills S-matrix*,
Phys. Rev. **D78** (2008) 125005, [`0807.4097`].

[85] 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.

[86] A. P. Hodges and S. Huggett, *TWISTOR DIAGRAMS*,
Surveys High Energ. Phys. **1** (1980) 333–353.

[87] L. J. Mason, *Twistor actions for non-self-dual fields: A Derivation of twistor-string theory*, *JHEP* **10** (2005) 009, [`hep-th/0507269`].

[88] R. Boels, L. J. Mason and D. Skinner, *Supersymmetric Gauge Theories in Twistor Space*, *JHEP* **02** (2007) 014, [`hep-th/0604040`].

[89] M. Abou-Zeid, C. M. Hull and L. J. Mason, *Einstein Supergravity and New Twistor String Theories*, *Commun. Math. Phys.* **282** (2008) 519–573, [`hep-th/0606272`].

[90] R. Boels, L. J. Mason and D. Skinner, *From twistor actions to MHV diagrams*, *Phys. Lett.* **B648** (2007) 90–96, [`hep-th/0702035`].

[91] N. Arkani-Hamed, F. Cachazo, C. Cheung and J. Kaplan, *The S-Matrix in Twistor Space*, *JHEP* **03** (2010) 110, [`0903.2110`].

[92] N. Arkani-Hamed, F. Cachazo, C. Cheung and J. Kaplan, *A Duality For The S Matrix*, *JHEP* **03** (2010) 020, [`0907.5418`].

[93] L. J. Mason and D. Skinner, *Dual Superconformal Invariance, Momentum Twistors and Grassmannians*, *JHEP* **11** (2009) 045, [`0909.0250`].

[94] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. B. Goncharov, A. Postnikov and J. Trnka, *Grassmannian Geometry of Scattering Amplitudes*. Cambridge University Press, 2016.

[95] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. Hodges and J. Trnka, *A Note on Polytopes for Scattering Amplitudes*, *JHEP* **04** (2012) 081, [`1012.6030`].

[96] N. Arkani-Hamed and J. Trnka, *The Amplituhedron*, *JHEP* **10** (2014) 030, [`1312.2007`].

[97] Z. Bern, J. J. M. Carrasco and H. Johansson, *New Relations for Gauge-Theory Amplitudes*, *Phys. Rev.* **D78** (2008) 085011, [`0805.3993`].

[98] F. Cachazo, S. He and E. Y. Yuan, *Scattering of Massless Particles in Arbitrary Dimensions*, *Phys. Rev. Lett.* **113** (2014) 171601, [`1307.2199`].

[99] F. Cachazo, S. He and E. Y. Yuan, *Scattering of Massless Particles: Scalars, Gluons and Gravitons*, *JHEP* **07** (2014) 033, [1309.0885].

[100] Z. Bern, T. Dennen, Y.-t. Huang and M. Kiermaier, *Gravity as the Square of Gauge Theory*, *Phys. Rev.* **D82** (2010) 065003, [1004.0693].

[101] J. J. Carrasco and H. Johansson, *Five-Point Amplitudes in N=4 Super-Yang-Mills Theory and N=8 Supergravity*, *Phys. Rev.* **D85** (2012) 025006, [1106.4711].

[102] J. J. M. Carrasco, M. Chiodaroli, M. Günaydin and R. Roiban, *One-loop four-point amplitudes in pure and matter-coupled N <= 4 supergravity*, *JHEP* **03** (2013) 056, [1212.1146].

[103] N. E. J. Bjerrum-Bohr, T. Dennen, R. Monteiro and D. O'Connell, *Integrand Oxidation and One-Loop Colour-Dual Numerators in N=4 Gauge Theory*, *JHEP* **07** (2013) 092, [1303.2913].

[104] R. H. Boels, R. S. Isermann, R. Monteiro and D. O'Connell, *Colour-Kinematics Duality for One-Loop Rational Amplitudes*, *JHEP* **04** (2013) 107, [1301.4165].

[105] M. Chiodaroli, Q. Jin and R. Roiban, *Color/kinematics duality for general abelian orbifolds of N=4 super Yang-Mills theory*, *JHEP* **01** (2014) 152, [1311.3600].

[106] Z. Bern, S. Davies, T. Dennen, Y.-t. Huang and J. Nohle, *Color-Kinematics Duality for Pure Yang-Mills and Gravity at One and Two Loops*, *Phys. Rev.* **D92** (2015) 045041, [1303.6605].

[107] J. Nohle, *Color-Kinematics Duality in One-Loop Four-Gluon Amplitudes with Matter*, *Phys. Rev.* **D90** (2014) 025020, [1309.7416].

[108] L. J. Dixon, *Calculating scattering amplitudes efficiently*, in *QCD and beyond. Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI-95, Boulder, USA, June 4-30, 1995*, pp. 539–584, 1996. hep-ph/9601359.

[109] J. J. M. Carrasco, *Gauge and Gravity Amplitude Relations*, 1506.00974.

[110] V. Del Duca, L. J. Dixon and F. Maltoni, *New color decompositions for gauge amplitudes at tree and loop level*, *Nucl. Phys.* **B571** (2000) 51–70, [hep-ph/9910563].

[111] R. Kleiss and H. Kuijf, *Multi - Gluon Cross-sections and Five Jet Production at Hadron Colliders*, *Nucl. Phys.* **B312** (1989) 616–644.

[112] C. Cheung, C.-H. Shen and C. Wen, *Unifying Relations for Scattering Amplitudes*, 1705.03025.

[113] Z. Bern, L. J. Dixon, D. C. Dunbar, M. Perelstein and J. S. Rozowsky, *On the relationship between Yang-Mills theory and gravity and its implication for ultraviolet divergences*, *Nucl. Phys.* **B530** (1998) 401–456, [hep-th/9802162].

[114] R. Monteiro, D. O'Connell and C. D. White, *Black holes and the double copy*, *JHEP* **12** (2014) 056, [1410.0239].

[115] R. Monteiro, D. O'Connell and C. D. White, *Gravity as a double copy of gauge theory: from amplitudes to black holes*, *Int. J. Mod. Phys.* **D24** (2015) 1542008.

[116] A. K. Ridgway and M. B. Wise, *Static Spherically Symmetric Kerr-Schild Metrics and Implications for the Classical Double Copy*, *Phys. Rev.* **D94** (2016) 044023, [1512.02243].

[117] A. Luna, R. Monteiro, I. Nicholson, D. O'Connell and C. D. White, *The double copy: Bremsstrahlung and accelerating black holes*, *JHEP* **06** (2016) 023, [1603.05737].

[118] W. D. Goldberger and A. K. Ridgway, *Radiation and the classical double copy for color charges*, *Phys. Rev.* **D95** (2017) 125010, [1611.03493].

[119] W. D. Goldberger, S. G. Prabhu and J. O. Thompson, *Classical gluon and graviton radiation from the bi-adjoint scalar double copy*, 1705.09263.

[120] A. Luna, R. Monteiro, D. O'Connell and C. D. White, *The classical double copy for Taub–NUT spacetime*, *Phys. Lett.* **B750** (2015) 272–277, [1507.01869].

[121] T. Adamo, E. Casali, L. Mason and S. Nekovar, *Scattering on plane waves and the double copy*, 1706.08925.

[122] D. G. Boulware and L. S. Brown, *Tree Graphs and Classical Fields*, *Phys. Rev.* **172** (1968) 1628–1631.

[123] M. J. Duff, *Quantum Tree Graphs and the Schwarzschild Solution*, *Phys. Rev.* **D7** (1973) 2317–2326.

[124] R. Monteiro and D. O'Connell, *The Kinematic Algebra From the Self-Dual Sector*, *JHEP* **07** (2011) 007, [1105.2565].

[125] C. Cheung and C.-H. Shen, *Symmetry and Action for Flavor-Kinematics Duality*, *Phys. Rev. Lett.* **118** (2017) 121601, [1612.00868].

[126] C. Cheung and G. N. Remmen, *Hidden Simplicity of the Gravity Action*, 1705.00626.

[127] R. Britto, *Loop Amplitudes in Gauge Theories: Modern Analytic Approaches*, *J. Phys.* **A44** (2011) 454006, [1012.4493].

[128] L. J. Dixon, *Scattering amplitudes: the most perfect microscopic structures in the universe*, *J. Phys.* **A44** (2011) 454001, [1105.0771].

[129] L. J. Dixon, *A brief introduction to modern amplitude methods*, in *Proceedings, 2012 European School of High-Energy Physics (ESHEP 2012): La Pommeraye, Anjou, France, June 06-19, 2012*, pp. 31–67, 2014. 1310.5353. DOI.

[130] H. Elvang and Y.-t. Huang, *Scattering Amplitudes in Gauge Theory and Gravity, Cambridge University Press* (2015) , [1308.1697].

[131] J. M. Henn and J. C. Plefka, *Scattering Amplitudes in Gauge Theories*, *Lect. Notes Phys.* **883** (2014) pp.1–195.

[132] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.