Perturbative Quantum Gravity from Gauge Theory

Zvi Bern

Department of Physics and Astronomy
UCLA, Los Angeles, CA
90095-1547, USA

Abstract

In this talk we give an overview of some recent developments in gauge and gravity theories, focusing on a new duality between color and kinematics. This duality allows us to construct gravity amplitudes by a simple replacement of color factors by kinematic numerator factors. Applications of these ideas for determining the ultraviolet compatibility of supersymmetric versions of Einstein gravity with quantum mechanics are explained.

1 Overview

Recent years have seen remarkable progress in understanding scattering processes of elementary particles in gauge and gravity theories, both for phenomenological and theoretical purposes. In this talk, we will focus on recent theoretical progress in quantum gravity. In particular, we will describe a surprising relation between gauge theory—used to describe nuclear forces—and gravity theories. For nearly 30 years, physicists have been convinced that point-like theories of gravity along the lines of Einstein’s theory are incompatible with quantum field theory because they lead to ultraviolet infinities. These infinites in turn lead to a loss of predictive power at ultra-high energies. Here we will describe how the new ideas relating gauge and gravity theories make it possible to challenge these beliefs by giving us the ability to carry out the required calculations.

Although gravity and gauge theories have obvious superficial similarities their detailed dynamics is rather different. Nevertheless we know from the celebrated AdS/CFT correspondence [1] between gauge and gravity theories that there is an equivalence between the weak coupling in one theory and strong coupling in the other. Here we will describe a different connection, purely at weak coupling, showing that in a precise sense gravity is a double copy of gauge theory. We write this schematically as

\[ \text{gravity} \sim (\text{gauge theory}) \times (\text{gauge theory}). \]  

This was first understood at tree level over 25 years ago using string theory [2], but today we have a much simpler description [3], allowing for a straightforward extension to loop level [4].

This new understanding of gravity has allowed us to probe the ultraviolet properties of gravity theories via explicit calculations at a level deeper than has been possible previously [5, 6, 7, 8]. Conventional wisdom holds that it is impossible to construct point-like ultraviolet finite quantum field theories of gravity (see e.g. ref. [9]).
This has been taken as a sign of a fundamental incompatibility between quantum field theory and gravity. Indeed, simple power-counting arguments show the difficulty of doing so. In a classic paper, ’t Hooft and Veltman demonstrated that gravity coupled to matter generically diverges at the first quantum loop order in four dimensions [10, 11]. Due to the dimensionful nature of the coupling, the divergences cannot be absorbed by a redefinition of the original parameters of the Lagrangian, rendering the theory non-renormalizable. Pure Einstein gravity does not possess a one-loop divergence [10, 12]. The two-loop divergence of pure Einstein gravity was established by Goroff and Sagnotti and by van de Ven through direct computation [13, 14]. Unfortunately, supersymmetry offers a mechanism for delaying the onset of divergences in gravity theories. No supergravity theory can diverge until at least three loops [9]. However, supersymmetry alone cannot eliminate the ultraviolet divergences in gravity theories because of the increasingly worse divergences at each loop order in gravity theories. This leads to the general question for supergravity theories of whether a given potential divergence identified by power counting and symmetry arguments alone is actually present. Here we will explain how the double-copy property (1) is helping us to resolve this question.

2 The on-shell philosophy

In recent years there has been a fundamental shift in how we view scattering amplitudes. In the traditional Feynman diagram approach one starts from an off-shell (i.e. one where states do not satisfy the Einstein relation $p^2 = m^2$) Lagrangian and constructs Feynman diagrams according to a set of rules. The diagrams encode algebraic expressions describing the scattering process. The diagrams depend on the gauge and field variable choices. Gauge invariance is restored only at the end of a computation, when one puts all external states on shell and all pieces are added together. The lack of gauge invariance for the individual diagrams can lead to enormously complicated expressions, which simplify only after a nontrivial effort to combine terms. In numerical approaches, it also exacerbates numerical instabilities.

In contrast, on-shell methods construct new amplitudes directly from simpler gauge-invariant on-shell amplitudes. Since the simpler amplitudes are already gauge invariant they can be greatly simplified before being used in the construction of more complex amplitudes. The two basic on-shell methods are on-shell recursion [15] and the unitarity method [16]. For studying multiloop gravity, the current method of choice is the unitarity method. This method was originally developed in the context of one-loop supersymmetric amplitudes [16], but with further refinements [17, 18, 19, 20, 21, 6, 22, 23], it offers a powerful formalism for any massless theory at any loop order, including non-planar contributions. This method has been reviewed numerous times [24, 25, 26], so here we give only a brief outline.

Unitarity has been a basic principle in quantum field theory since its inception. For a description of unitarity during the 1960’s see ref. [27]. However, a variety of difficulties prevented its widespread use as a means of constructing amplitudes, especially after the rise of gauge theories in the 1970’s. These difficulties include non-convergence of dispersion relations and its inapplicability to massless particles. It was also unclear how
one could fully reconstruct loop amplitudes beyond four points from their unitarity cuts. The modern unitarity method overcomes these difficulties, allowing for the complete construction of loop amplitudes at any loop order. It does so by avoiding dispersion relations, and instead using the existence of an underlying covariant Feynman diagram representations to fully reconstruct amplitudes. By construction the obtained Feynman-like integrands have the correct analytic properties in all channels.

Over the years there have been a number of important refinements to the unitarity method [16]. Generalized unitarity [27] (where multiple internal lines are placed on shell, subdividing a loop amplitude into more than two pieces) was successfully applied in ref. [18] as a means for greatly simplifying loop calculations. An important more recent development is the use of complex momenta [13] by Britto, Cachazo and Feng [19], leading to the realization that at one loop in four dimensions, quadruple cuts directly determine the coefficients of all box integrals by freezing the loop integration. Powerful new methods for dealing with triangle and bubble integrals at one loop, as well as rational terms have also been developed [17, 21, 28, 23, 29]. (These have been described in other recent reviews [30].) At higher loops, efficient means of constructing the integrands of amplitudes, including non-planar contributions, have also been devised [20, 6, 22, 31].

Although the unitarity method applies just as well to supersymmetric and non-supersymmetric theories, it is usually much simpler to deal with the supersymmetric cases because they have a simpler analytic structure. Indeed, the original application of the unitarity method was to construct one-loop supersymmetric amplitudes with arbitrary numbers of external legs [16].

3 Comparing Gravity to Gauge Theory

We start by comparing gravity to gauge theory using off-shell methods. The Feynman rules are generated starting from the Einstein-Hilbert and Yang-Mills Lagrangians,

\[ \mathcal{L}_{YM} = -\frac{1}{4} F^{a}_{\mu \nu} F^{a \mu \nu}, \quad \mathcal{L}_{EH} = \frac{2}{\kappa^2} \sqrt{-g} R. \]  

(2)

From the viewpoint of Feynman diagrams, these two Lagrangians have rather different properties. With standard gauge choices gauge theories have three- and four-point interactions, while gravity has an infinite number of contact interactions. Perhaps more striking than the infinite number of interactions is the remarkably complexity of these interactions.

To be more concrete, consider the three-gluon vertex in Feynman gauge,

\[ V^{abc}_{3\mu,\nu,\sigma}(k_1, k_2, k_3) = g f^{abc} \left( (k_1 - k_2)_{\sigma} \eta_{\mu \nu} + \text{cyclic} \right), \]  

(3)

where \( g \) is the coupling, \( f^{abc} \) the usual group theory structure constants, \( \eta_{\mu \nu} \) the flat metric and the \( k_i \) the momenta of the vertex. This vertex is relatively simple. We may compare this to the three-graviton interaction in, for example, de Donder gauge,

\[ G_{3\mu \alpha, \nu \beta, \sigma \gamma}(k_1, k_2, k_3) = i \frac{\kappa}{2} \left[ -\frac{1}{2} k_1 \cdot k_2 \eta_{\mu \alpha} \eta_{\nu \beta} \eta_{\sigma \gamma} - \frac{1}{2} k_1 \eta_{\mu \alpha} k_2 \eta_{\nu \beta} \eta_{\sigma \gamma} + \cdots \right], \]  

(4)
where we have displayed two terms out of about 100. Here the coupling $\kappa$ is related to Newton’s constant by $\kappa^2 = 32\pi^2 G_N$. The precise form of the vertex depends on the gauge, but in any case the three vertex is a rather involved and unenlightening object. The complete expression can be found in refs. [32, 10].

Comparing the vertex in eq. (3) to the one in eq. (4), it certainly would appear that gravity is much more complicated than gauge theory. Moreover, there does not appear to be any simplicity or obvious relation between the gauge and gravity vertices. The former leads to complicated diagrams, but the latter appears hopelessly complicated. One can do somewhat better with special gauge choices and appropriate field redefinitions [14, 33], considerably simplifying the Feynman rules. Still, multiloop Feynman diagram calculations in (super) gravity are extremely difficult, and generally out of reach using even the most powerful supercomputers.

Now let us reconsider the same process but from the on-shell vantage point. If we take the three-graviton vertex in eq. (4) and dot the three legs with physical polarizations satisfying the physical state conditions, $k^2_i = 0$, $\varepsilon^\mu k^\mu_i = \varepsilon^\mu k^\nu_i = \varepsilon^\mu = 0$, we obtain a greatly simplified vertex,

$$G_3(k_1, k_2, k_3) = -i\kappa \varepsilon_1^\mu \varepsilon_2^\nu \varepsilon_3^\gamma (k_1)_\sigma \eta_{\mu\nu} + \text{cyclic} \left[(k_1)_\gamma \eta_{\alpha\beta} + \text{cyclic}\right]. \tag{5}$$

Remarkably, up to overall factors, this is just a double copy of the kinematic part of the on-shell Yang-Mills vertex,

$$V_3^{abc}(k_1, k_2, k_3) = 2\varepsilon_1^\mu \varepsilon_2^\nu \varepsilon_3^\sigma g f^{abc} (k_1)_\sigma \eta_{\mu\nu} + \text{cyclic}, \tag{6}$$

where the polarization vector satisfies $\varepsilon^\mu_i k^\mu_i = 0$. To make the comparison, we identify the graviton polarization tensor as a product of gluon polarization vectors, $\varepsilon^\mu_i = \varepsilon^\mu_i \times \varepsilon^\mu_i$. Similar considerations allow us to express all three-point vertices in supergravity as products of super-Yang-Mills vertices. Using BCFW recursion [15], these three vertices are sufficient to construct any tree-level gauge or gravity amplitude. The unitarity method then allows us to construct any loop amplitude.

Clearly, there is a rather striking relationship between gravity and gauge theory, but to make it visible we need to keep external states on shell. As we shall see below, the double-copy structure in eq. (5) is not accidental, but appears likely to extend to all loop orders. As such, it reflects a profound and important property of quantum gravity, pointing to unification of the two theories, perhaps along the lines of string theory.

Along these lines, we now discuss the recently discovered duality between color and kinematics [3, 4]. In general, we can write any $n$-point tree-level gauge-theory amplitude with all particles in the adjoint representation as,

$$A_n^{\text{tree}}(1, 2, 3, \ldots, n) = \sum_i \frac{n_i c_i}{\prod_{\alpha_i} \mu_{\alpha_i}^2}, \tag{7}$$

where the sum runs over the set of $n$-point $L$-loop diagrams with only cubic vertices. These include distinct permutations of external legs. We have suppressed factors of
the coupling constant for convenience. The product in the denominator runs over all propagators of each cubic diagram. The $c_i$ are the color factors obtained by dressing every three vertex with an $f^{abc} = i\sqrt{2}f^{abc}$ structure constant, and the $n_i$ are kinematic numerator factors depending on momenta, polarizations and spinors. The form (7) can be obtained straightforwardly, for example, from Feynman diagrams, by representing all contact terms as inverse propagators in the kinematic numerators that cancel propagators. For supersymmetric amplitudes expressed in superspace, there will also be Grassmann parameters in the numerators.

The duality conjectured in ref. [3] requires there to exist such a transformation from any valid representation to one where the numerators satisfy equations in one-to-one correspondence with the Jacobi identity of the color factors,

$$c_i = c_j - c_k \Rightarrow n_i = n_j - n_k .$$ (8)

This duality is conjectured to hold to all multiplicity at tree level in a large variety of theories, including supersymmetric extensions of Yang-Mills theory. In fig. 1 we display the Jacobi relation at four points. The duality conjecture states there exists representations of the amplitude, such that the color factors and numerators of the diagrams satisfy the relations.

At tree level, a consequence of this duality is non-trivial relations between the color-ordered partial tree amplitudes of gauge theory [3, 34, 35]. The duality has also been studied in string theory [36, 37] and in terms of Lagrangians [38]. An alternative trace-based representation of the duality (8) was recently given in ref. [39], emphasizing the underlying group theoretic structure of the duality. In the self-dual case, underlying group theoretic structure has been made explicit [40].

Perhaps more remarkable than the duality itself is a related conjecture that once the gauge-theory amplitudes are arranged into a form satisfying the duality (8), corresponding gravity amplitudes can be obtained simply by replacing the $c_i$ color factor in eq. (7) with a second copy of a numerator factor $\tilde{n}_i$ [3, 4],

$$-iM^{\text{tree}}_n (1, 2, \ldots, n) = \sum \frac{n_i \tilde{n}_i}{\prod_{\alpha} p_{\alpha_i}^2} ,$$ (9)

The sum runs over the same set of diagrams with cubic vertices, as in eq. (7). This is expected to hold in a large class of gravity theories, including theories that are the low-energy limits of string theories. (As for the gauge-theory case, we suppress factors of the coupling constants.) At tree level, this double-copy property encodes what are known as
KLT relations between gravity and gauge-theory tree amplitudes [2]. The double-copy formula (9) has been proven via on-shell recursion [15] for pure gravity and for $\mathcal{N} = 8$ supergravity tree amplitudes, whenever the duality (8) holds in the corresponding gauge theories [38].

More recently, the above conjectures have been extended to loop level [4], so that at any loop order $L$,

$$
\mathcal{A}_n^{\text{loop}} = \sum_j \int \prod_{l=1}^L \frac{d^D p_l}{(2\pi)^D} \frac{1}{S_j \prod_{\alpha_j} p_{\alpha_j}^2}, \quad \mathcal{M}_n^{\text{loop}} = \sum_j \int \prod_{l=1}^L \frac{d^D p_l}{(2\pi)^D} \frac{1}{S_j \prod_{\alpha_j} p_{\alpha_j}^2},
$$

(10)

where $\mathcal{A}_n^{\text{loop}}$ and $\mathcal{M}_n^{\text{loop}}$ are $L$-loop gauge and gravity amplitudes. As before we removed factors of the coupling constants. The sums now run over all distinct $m$-point $L$-loop diagrams with cubic vertices. These include distinct permutations of external legs, and the $S_j$ are the symmetry factors of each diagram. As at tree level, at least one family of numerators ($n_j$ or $\tilde{n}_j$) for gravity must be constrained to satisfy the duality (8). (For pure gravity, extra projectors are needed to obtain loop-level amplitudes from the direct product of two pure Yang-Mills theories.) A three-loop example of a duality relation for numerators factors is displayed in fig. 2. For the duality (8) to hold, the duality relation for every propagator in all diagrams must be enforced.

This loop-level extension has been tested in the rather nontrivial case of three- and four-loop four-point amplitudes [4, 8] and two-loop five point amplitude [41] of $\mathcal{N} = 4$ super-Yang-Mills theory and $\mathcal{N} = 8$ supergravity. It has also been tested in one and two-loop gravity examples in cases with fewer supersymmetries than the maximum [42].

4 Ultraviolet properties of gravity

Today, our most powerful tool for explicitly determining the ultraviolet properties of gravity theories is the double-copy property used in tandem with the unitarity method, as described in a very recent paper [8]. The best theories to study are the maximally supersymmetric ones, because of their technical simplicity and because supersymmetry tends to mitigate ultraviolet divergences. $\mathcal{N} = 4$ super-Yang-Mills theory was proven to be ultraviolet finite in four dimensions long ago [43]. The ultraviolet behavior of
\( \mathcal{N} = 8 \) supergravity [44] in four dimension is, however, still under study. Recent reviews discussing the ultraviolet properties of \( \mathcal{N} = 8 \) supergravity in more detail are given in refs. [45].

Because of the complexity of explicit calculations, people normally resort to power counting arguments. These arguments assume that all symmetries and relevant properties are known and accounted for. For the case of \( \mathcal{N} = 8 \) supergravity we know that there are unexpected ultraviolet cancellations in the theory to all loop orders [46, 47], though it is still not clear if these are powerful enough to render the theory finite to all loop orders. (These cancellations are related to a well studied property of one-loop \( \mathcal{N} = 8 \) amplitudes: in four dimensions triangle and bubble integrals drop out of the amplitudes, when expressed in a basis of scalar integral in four dimensions [48].) Some hints also follow from string-theory dualities [49]. We also know that gravity loop amplitudes are much more closely tied to better behaved gauge-theory amplitudes than had been believed [4]. While these arguments do not offer a proof of finiteness, they do suggest that it would be wise to reexamine the ultraviolet properties of gravity theories. For other approaches to trying to make quantum field theories of gravity sensible in the ultraviolet see refs. [50].

Motivated by the hint of high-loop cancellations, explicit calculations were carried out in refs. [5, 6, 7, 8] to directly investigate the ultraviolet properties of \( \mathcal{N} = 8 \) supergravity. These calculations definitively rule out the expected potential three-loop divergence in four space-time dimensions. Although no potential divergence exists at four loops in four dimensions (because of an “accidental” cancellation similar to the one preventing a pure gravity divergence at one loop), direct calculation establishes that the four-loop four-point amplitude of \( \mathcal{N} = 8 \) supergravity has the same power counting in \( D \) dimensions as \( \mathcal{N} = 4 \) super-Yang-Mills theory (which is known to be finite in \( D = 4 \)).

The result of direct calculation [7, 8] is that the four-loop four-point amplitude of \( \mathcal{N} = 8 \) supergravity is of the form,

\[
M_{4\text{-loop}}^{4\text{-point}} \sim D^8 R^4 \times \text{loop integrals}
\]

where the \( D^8 R^4 \) factor corresponds to 16 powers of momentum in the numerators of the integrals coming out as external momentum. This factor is a shorthand for covariant derivatives acting on four Riemann tensors with their Lorentz indices contracted in an appropriate way. If we assume that no further ultraviolet cancellations exist, and that no further powers of loop momenta can come out of the integrals as external momenta as the loop order increases, simple power counting shows that in four dimensions the first divergence would occur at seven loops.

This is in line with recent comprehensive studies of the potential divergences in \( \mathcal{N} = 8 \) supergravity [51, 52, 53, 54], showing that no divergence is compatible with the known symmetries until seven loops. Based on these studies, a consensus has formed that symmetry constraints alone cannot prevent divergences in four space-time dimensions starting at seven loops and that the theory will likely diverge at this loop order. There is, however, a more optimistic view [55]. (We note that the previously claimed delay until nine loops of potential ultraviolet divergences in \( \mathcal{N} = 8 \) supergravity [56] has now been retracted [52].)

Is it possible that there are further symmetries or structures that prevent the widely
expected seven loop divergences? Powercounting arguments using symmetries to rule out potential divergences can, of course, never prove the existence of divergences, only that protection against divergences holds to a certain level; if a symmetry or structure is missed then it may turn out the bound is too loose. More generally, the only way we can be certain that the coefficient of a potential divergence respecting the known symmetries is non-zero is to carry out the explicit calculation to show that the numerical value is nonzero.

Today, even with all the advances, it is not yet practical to carry out a seven-loop computation. However, a simple way to lower the loop order in which a given potential divergence can occur is to work in higher space-time dimensions higher dimensions. By increasing the dimension, \( \mathcal{N} = 4 \) super-Yang-Mills is no longer ultraviolet finite, allowing this theory to be used as a playground for sharpening our understanding of divergences in maximally supersymmetric theories \([6, 57, 22]\). Explicit calculations \([58, 20, 5, 6, 7, 22]\) show that at least for four-point amplitudes through four loops, both \( \mathcal{N} = 8 \) supergravity and \( \mathcal{N} = 4 \) super-Yang-Mills theory are ultraviolet finite for

\[
D < \frac{6}{L} + 4 \quad (L > 1),
\]

where \( D \) is the dimension of space-time and \( L \) the loop order. (The case of one loop, \( L = 1 \), is special, with the amplitudes finite for \( D < 8 \), not \( D < 10 \).) For \( \mathcal{N} = 4 \) super-Yang-Mills this bound was proposed in ref. \([58]\) and has been confirmed in ref. \([59]\) using superspace techniques. Explicit computations summarized below demonstrate this bound is saturated in \( \mathcal{N} = 4 \) super-Yang-Mills theory through at least four loops \([60, 58, 46]\). For \( \mathcal{N} = 8 \) supergravity we know that the bound (12) is saturated through four loops \([58, 5, 6, 7, 8]\).

5 Outlook

In this talk we described a surprising relation between gravity and gauge theories, stemming from a gauge-theory duality between color and kinematics \([3, 4]\). Although the duality still has the status of a conjecture at loop level, we can exploit it to streamline loop computations based on the unitarity method. These types of computations have been successfully used to probe the ultraviolet properties of supergravity theories.

The current consensus in the community is that the standard symmetries of \( \mathcal{N} = 8 \) supergravity cannot protect the theory against divergences, starting at seven loops \([51, 52]\) (though there is at least one contrary opinion \([55]\)). If divergence do appear at seven loop, then we should see indications starting at five loops, albeit in higher space-time dimensions. To test this, it would be of crucial importance to directly determine the ultraviolet properties of \( \mathcal{N} = 8 \) supergravity as a function of dimension at five loops. If this calculation can be completed, it should greatly clarify the ultraviolet behavior of \( \mathcal{N} = 8 \) supergravity in four dimensions, checking the hypothesis that it is an ultraviolet finite theory. As recently discussed in some detail in ref. \([8]\), the duality between color and kinematic numerators and the associated double-copy property of gravity offers a
promising approach to solve this problem. We can look forward to many new exciting results in the coming years based on these developments.

Acknowledgments

We thank S. Davies, L. J. Dixon, J. J. M. Carrasco, T. Dennen, H. Johansson, Y.-t. Huang, H. Ita, D. A. Kosower, and R. Roiban for many enlightening discussions and collaboration on work described in this lecture. This work was supported by the US Department of Energy under contract DE-FD03-91ER40662. I wish to thank the organizers for inviting me to such a wonderful conference.

References


