La Rivista del Nuovo Cimento (2020) 43:97–186
https://doi.org/10.1007/s40766-020-00003-6

REVIEW PAPER

Gravity as the square of gauge theory: a review

L. Borsten[1]

Received: 18 November 2019 / Accepted: 12 January 2020 / Published online: 11 March 2020

Abstract

We briefly review the manifold incarnations and facets of the notion that gravity, in certain regards, can be re-conceived as the product of two gauge theories. We begin with a short history of the relationship between gauge and gravity theories in the context of "gravity $=$ gauge $\times$ gauge". This is followed by modern approaches to gauge/gravity scattering amplitude relations, focussing on the Bern–Carrasco–Johansson colour-kinematic duality and double-copy construction of gravitational amplitudes from gauge amplitudes. This includes a partial characterisation of what gravity theories admit a gauge theory squared origin. We then consider classical and off-shell perspectives on "gravity $=$ gauge $\times$ gauge". First we review field theoretic approaches to understanding the colour-kinematic duality and the double-copy prescription, including kinematic algebras and colour-kinematic duality and double-copy manifesting Lagrangians. We then consider classical double-copy-like methods for constructing gravitational solutions from gauge theory solutions. To close, we consider a purely field theoretic take on "gravity $=$ gauge $\times$ gauge". This framework centres on a convolutive product of gauge theories, at the level of spacetime fields themselves, that yields gravitational theories.

Contents

L. Borsten
leron@stp.dias.ie

1 School of Theoretical Physics, Dublin Institute for Advanced Studies, 10 Burlington Road, Dublin 4, Ireland

1 Introduction

The early part of the twentieth-century witnessed two revolutions that laid out the foundations of physics today: quantum theory and Einstein's general theory of relativity. While each pillar has been tremendously successful in its respective domain, they seemingly stand apart from a number of perspectives. At the sharpest end of the apparent contradictions, general relativity and quantum theory clash with seemingly paradoxical consequences [1,2], leaving our account of Nature fragmented and incomplete. This is the problem of quantum gravity. While there are a number of promising paths to quantum gravity, string/M-theory being the most relevant in the context of this review, its ultimate resolution remains a challenge for twenty-first-century physics. Even leaving these most dramatic schisms aside, general relativity and quantum theory paint curiously distinct pictures of the known fundamental forces. The electroweak and strong forces of the Standard Model of Particle Physics are described by quantised Yang–Mills gauge theories that play out on a fixed spacetime stage. On the other hand, gravity is most economically thought of as the curvature of spacetime itself, which is elevated to a dynamical degree of freedom in its own right. Moreover, the gauge theories underpinning the Standard Model are renormalisable, whereas the straightforward perturbative quantisation of general relativity, with or without matter, is plagued by uncontrollable divergences [3–7].

From this point of view it would seem that gravity cuts a lonely figure. However, the diffeomorphism invariance of general relativity makes it in a sense the "gauge theory" *par excellence*. Thought of this way perhaps gravity and gauge theory are not so distant after all. Indeed, the hope that gravity can be understood in terms of gauge theory, or vice versa, has been a reoccurring theme in the annuls of theoretical physics. It has taken numerous guises, many of which are inter-related. The earliest example is provided by Kaluza–Klein theory [8–10]; general relativity in $D = 5$ spacetime dimensions compactified on a circle S^1 includes in its massless spectrum a $\mathrm{Un}(1)$ gauge field.[1] Electromagnetism is seen to derive from pure geometry and gravity. Turning this relation around, the first and perhaps most logically transparent approach is to gauge the Lorentz, Poincaré or de Sitter symmetries [12–16]. Indeed, general relativity in its first-order form, which given the need to conveniently couple to fermions is not merely a cosmetic rewriting, has a manifest local Lorentz symmetry.

[1] For a review of Kaluza–Klein theory see [11].

The spin-connection $w^a{}_b$ is the corresponding gauge field and its field strength is nothing but the Riemann tensor, $R^a{}_b = dw^a{}_b + w^a{}_c \wedge w^c{}_b$. This would seem to make the relationship to Yang–Mills theory manifest, but more care is needed. In fact, the necessary extra ingredients make the differences between the Yang–Mills theory of local Poincaré symmetry and general relativity quite clear. Gauging the Poincaré group[2] yields a covariant derivative $D = d + A$ with connection

$$A = e^a P_a + w^a{}_b M^b{}_a \tag{1}$$

where P_a and $M^b{}_a$ are the translation and Lorentz generators, respectively, and we have suggestively labelled the corresponding gauge fields by e^a and $w^a{}_b$. The Poincaré algebra, $[M, P] \sim P$, $[M, M] \sim M$, $[P, P] = 0$, implies the field strength takes the form

$$F = [D, D] = T^a P_a + R^a{}_b M^b{}_a \tag{2}$$

where

$$R^a{}_b = dw^a{}_b + w^a{}_c \wedge w^c{}_b, \quad T^a = de^a + w^a{}_b \wedge e^b. \tag{3}$$

Hence, identifying e^a and $w^a{}_b$ with the frame field and spin-connection, respectively, T^a and $R^a{}_b$ correspond to the torsion and Riemann tensors. However, on the one hand the Yang–Mills Lagrangian $\propto \mathrm{tr}(F \wedge \star F)$ obviously does not yield the Einstein equation (and, besides, the Hodge dual requires a metric, somewhat undermining the whole approach) and, on the other hand, the Einstein–Hilbert Lagrangian $\propto \sqrt{-g}R$ is not gauged Poincaré invariant (and again requires a metric). Both deficiencies can be rectified using the Plebanski Lagrangian $\propto e^a \wedge e^b \wedge R^{cd}\epsilon_{abcd}$, which is invariant under local Poincaré transformations *up to a term proportional to the torsion* T^a. So, we have a local Poincaré invariant action that yields the Einstein equation *if* we impose that the torsion vanishes, $T^a = 0$. With this condition in place, we may identify the local Pioncaré translation parameter $\epsilon^a(x)$ with a vector ξ, via $\epsilon^a = i_\xi e^a$, so that the local Poincaré transformation of e^a becomes a Lie derivative with respect to ξ together with a shifted local Lorentz rotation,

$$\delta e^a = D\epsilon^a + \alpha^a{}_b e^b = \mathcal{L}_\xi e^a - i_\xi T^a + \left(\alpha^a{}_b - i_\xi w^a{}_b\right) e^b \underset{T^a \to 0}{\longrightarrow} \mathcal{L}_\xi e^a + \alpha^a_{(\xi)}{}_b e^b \tag{4}$$

So far, so good. However, while the $T^a = 0$ constraint is consistent with the equations of motion, one should still object that it is not itself gauged Poincaré invariant, $\delta T^a|_{T^a \to 0} = \left(R^a{}_b \epsilon^b + \alpha^a{}_b T^b\right)|_{T^a \to 0} = R^a{}_b \epsilon^b$. In this sense, general relativity is strictly *not* given by gauging the Poincaré group. Nonetheless, the procedure of gauging spacetime symmetries and then identifying appropriate constraints is remarkably

[2] To be absolutely clear, here we are considering the Yang–Mills theory of the Poincaré group in the strict sense that it is treated as an internal symmetry, in direct analogy to, say, the SU(3) of the Standard Model. By "gauged Poincaré" we really do mean a principal bundle $P(M, G)$, for G the Poincaré group. Of course, the actual role of the global Poincaré group in a relativistic quantum field theory on $\mathbb{R}^{1,d}$ is as a spacetime symmetry. We can also promote this spacetime Poincaré group to be local, as was done in [13]. In this case, the local translations and rotations yield independent general coordinate and local Lorentz transformations. Demanding invariance then leads one, almost inevitably, to general relativity (with matter induced torsion) without invoking, *a priori*, any Riemannian geometry [13]. While certainly elegant, this seems to us to be gilding the lily somewhat.

effective and, for example, has been applied to construct Poincaré, anti-de Sitter and conformal supergravity theories [15,17,18], a highly non-trivial task when tackled conventionally. Moreover, the explicit imposition of constraints can be replaced by a spontaneous symmetry breaking mechanism of a larger group [19]. Such successes notwithstanding, a metric on an n-dimensional Riemannian manifold can be regarded as a section of the associated frame bundle with fibres $\mathrm{GL}(n)/\mathrm{O}(n)$, so is closer in spirit to a gauged non-linear sigma model equipped with a soldering form, rather than a conventional Yang–Mills theory of the Poincaré group. This takes us rather beyond the direct connection to Yang–Mills gauge theory and certainly the scope of this review.

However, this is not the only way one might think to relate gauge theory to gravity. In particular, the holographic principle [20,21], realised concretely through the Anti-de Sitter/Conformal Field Theory (AdS/CFT) correspondence [22–24], represents a subtler and ultimately deeper gauge/gravity relation. In fact, the idea that the spin-two graviton is a composite of two spin-one gauge bosons can be taken as the starting point of a heuristic route to AdS/CFT, without appealing *a priori* to string theory[3] [27,28]. Let us expand on this point of view briefly, as it is instructive and serves to highlight the crucial differences between the AdS/CFT gauge/gravity duality and "gravity $=$ gauge $\times$ gauge". It has been long-ago suggested, more than once, that the spin-two graviton is not elementary, but rather a bound state appearing in some renormalisable quantum field theory [29–32]. On the basis of representation theory alone it is not unreasonable to suppose that a spin-2 graviton (coupled to a scalar) might be composed as the symmetric tensor product of two spin-1 particles (or even four spin-1/2 particles [29]). The problem with this idea is that the Weinberg–Witten theorem [33] would seem to rule out massless composite particles of spin greater than one in any quantum field theory, under the assumption that there exists a conserved Lorentz covariant energy-momentum tensor. However, along with requiring a stress tensor, there is another hidden assumption so seemingly innocuous that it almost does not bear mentioning; the composite graviton lives in the same spacetime as its elementary constituents. But this is precisely what the holographic principle violates: a theory in $(D+1)$ spacetime dimensions is the "holographic image" of a theory in D dimensions. The Weinberg–Witten theorem does *not* exclude the possibility that a graviton propagating in a $(D+1)$-dimensional bulk spacetime is equivalently described in terms of a gauge theory living on a D-dimensional boundary. Of course, things are not quite so simple. First, in order to capture the physics of a local $(D+1)$-dimensional theory the boundary theory needs an additional "dimension"—a parameter with respect to which the physical observables are local. This is provided by the energy scale μ; the couplings are governed by the renormalisation group flow equations, which are local functions of μ. Second, this scale should be macroscopic. Coupled with the expectation that the boundary theory should be highly quantum, since a perturbative gauge theory does not look like classical gravity, this suggests that the boundary theory is strongly-coupled over a large range of energies. For infinite bulk radial dimension the boundary theory should therefore be conformal. Supersymmetry can be employed to retain control in the strongly-coupled regime, in which case the gauge

[3] Although one is inevitably led back to string theory [25], as with hindsight might have been anticipated from 't Hooft's work on the planar limit of Yang–Mills theory [26].

theory is superconformal. Finally, to account for the extra classical functional degrees of freedom enjoyed by fields propagating in $D+1$ dimensions with respect to those in D dimensions, a large gauge group limit is invoked. For example, the equivalence of Yang–Mills theory, with $\mathrm{SU}(N)$ colour group and gauge coupling g, with classical gravity is valid for $N \gg \lambda \gg 1$, where we have introduced the ‘t Hooft coupling $\lambda = g^2 N$. The best understood case, which is by now extremely well-tested, has type IIB string theory on an asymptotically $\mathrm{AdS}_5 \times S^5$ bulk spacetime, with the maximally supersymmetric $D=4$ Yang–Mills theory on the AdS_5 boundary. It is widely believed to be a complete duality, also holding for finite N and λ, although this is much harder to test. Regardless, even in its most conservative form, $N \to \infty$, the AdS/CFT correspondence provides a remarkable gauge/gravity relation, with profound implications and myriad applications.

In the present contribution, our concern is a third (at least naively) independent relationship between gauge theory and gravity: the “gravity = gauge × gauge” paradigm. While it also takes as its starting point the idea that spin-2 can be built from spin-1, it is quite different from a number of perspectives. First, it is rather generic, not requiring anything like large N, strong/weak-coupling-duality, supersymmetry or the holographic principle. It does however require other hidden properties, which are shared by a very broad class of gauge theories. We shall come to those momentarily. On the other hand, it is rather more limited in the sense that it is not a complete duality, but rather a growing set of compelling relations. To be more precise, we simply do not yet know just how general or deep it is. In this review, we shall describe the various and connected perspectives on “gravity = gauge × gauge” and explore just how far it can be taken. Let us state at the outset that we do not have a definitive answer - the jury is still out. However, the proliferation of surprising, illuminating and powerful insights uncovered thus far compels further serious consideration, its rather radical nature notwithstanding.

The heuristic picture to have in mind is that one can regard the product of two gauge potentials as a gravitational theory as described by the suggestive, but for the moment purely illustrative, equation:

$$\text{“}A_\mu \otimes \tilde{A}_\nu = g_{\mu\nu} \oplus B_{\mu\nu} \oplus \varphi\text{”}. \tag{5}$$

Here, A_μ and $\tilde{A}_\nu$ are the gauge potentials of two distinct Yang–Mills theories, which we will refer to as the left (no tilde) and right (tilde) theories, or factors, respectively. They can have arbitrary and independent non-Abelian gauge groups G and $\tilde{G}$. Their “product” yields a metric, $g_{\mu\nu}$, an Abelian 2-form gauge potential $B_{\mu\nu}$ and a scalar field φ. This equation has meaning if we interpret it as the tensor product of the corresponding spacetime little group representations, but going beyond this requires a rather more subtle approach. Indeed, there are a number of good reasons, Weinberg–Witten aside, to suspect it cannot be well-defined. For one the right hand side should be covariant with respect to general coordinate transformations whereas the left hand side transforms locally under two arbitrary finite-dimensional Lie groups. Nonetheless, the idea represented by (5) has proven itself incredibly powerful, particularly in the context of scattering amplitudes, motivating a reappraisal of these apparent obstructions.

Given the lessons of Weinberg–Witten (forbidding composite particles) and AdS/CFT (relying on a holographic dimension), how could such a proposal work? The first substantial clues came, just as in the case of AdS/CFT, from string theory in the guise of the Kawai–Lewellen–Tye (KLT) relations [34], which connect the tree-level scattering amplitudes of closed strings to sums of products of open string amplitudes. While highly non-trivial, the intuition underpinning these relations is quite clear.[4] First, the spectra of closed strings is given by the tensor product of those corresponding to left and right moving open strings. Since the low energy effective field theory limits of closed and open superstrings are given by supergravity and super Yang–Mills theory, respectively, graviton states arise as the tensor product of the gluon states.

For instance, the massless sector of type I superstrings in $D = 10$ is given by super Yang–Mills theory [36] with gluons and gluini, that is adjoint-valued Majorana–Weyl (MV) spinors, in the $\mathbf{8}_v$ and $\mathbf{8}_s$, respectively, of the (double cover of) the spacetime little group Spin(8). The product of type I superstrings with opposing chiralities yields type IIA superstrings, with massless spectrum,

$\otimes$	$\mathbf{8}_v$	$\mathbf{8}_c$
$\mathbf{8}_v$	$\mathbf{35}$ (graviton) $+$ $\mathbf{28}$ (KR 2-form) $+$ $\mathbf{1}$ (dilaton)	$\mathbf{56}_s$ (gravitino) $+$ $\mathbf{8}_s$ (MW spinor)
$\mathbf{8}_s$	$\mathbf{56}_c$ (gravitino) $+$ $\mathbf{8}_c$ (MW spinor)	$\mathbf{56}_v$ (RR 3-form) $+$ $\mathbf{8}_v$ (RR 1-form)

(6)

In particular, we see that the on-shell *gluon* $\otimes$ *gluon* sector[5] yields an on-shell graviton $\mathbf{35}$, the on-shell Kalb–Ramond (KR) 2-form $\mathbf{28}$, and the dilaton $\mathbf{1}$, just as anticipated by the heuristic formula (5). Of course, this is just a special case of $V \otimes V \cong \mathrm{Sym}^2(V)_0 \oplus \wedge^2(V) \oplus \mathbb{R}$, where $V \cong \mathbb{R}^{D-2}$ is the vector representation of $\mathrm{SO}(D-2)$ and $\mathrm{Sym}^2(V)_0$ denotes the trace free symmetric product. This is the universal sector of the product all conventional (super) Yang–Mills theories for any dimension D and any number of allowed supersymmetries $\mathcal{N}$. It is therefore useful to give it a name. It is sometimes referred to as "$\mathcal{N} = 0$ supergravity' and we shall adopt this convention. In addition to the universal bosonic sector, we have the fermions from the *gluon* $\otimes$ *gluino* + *gluino* $\otimes$ *gluon* sector consisting of two gravitini, $\mathbf{56}_s$ and $\mathbf{56}_c$, and two MV spinors, $\mathbf{8}_c$ and $\mathbf{8}_s$. The appearance of two gravitini with opposing chiralities implies that there are $\mathcal{N} = 2 = 1 + 1$ local supersymmetries corresponding to type IIA supergravity. This reveals our second universal property apparent just at the level of representation theory; the global left and right supersymmetries sum in the product to give local supersymmetries.[6] Finally, in the *gluino* $\otimes$ *gluino* sector we find the R-R 1-form $\mathbf{8}_v$

4 An excellent account can be found in §7 of the classic reference [35].

5 In the stringy nomenclature we have the NS–NS ($\mathbf{8}_v \times \mathbf{8}_v$), NS-R ($\mathbf{8}_v \times \mathbf{8}_c$), R–NS ($\mathbf{8}_s \times \mathbf{8}_v$), and R–R ($\mathbf{8}_s \times \mathbf{8}_c$) sectors, where NS and R refer to Nevau–Schwarz and Ramond.

6 On the one hand, this is exactly what one would expect from our stringy intuition based on world-sheet supersymmetry, which is treated democratically in both the open string factors and their closed string product. It is, however, a little counter-intuitive from the spacetime "gravity = gauge × gauge" perspective, where we take internal tensor products of spacetime little group representations, but external tensor products of the supersymmetry transformations, despite the fact that they anti-commute to give translations. It all works out, as we shall see.

and 3-form $\mathbf{56}_v$ Abelian gauge fields. This is just the Clifford expansion of the spinor tensor products.

The matching of closed string spectra with the products of open string spectra is, of course, just the minimal requirement for anything like the KLT relations to work, but also allows for (or, rather, follows from) the construction of closed string vertex operators in terms of those left and right moving open strings, $V_{\text{closed}}(2p, \tau) \sim \int d\sigma V_{\text{open}}(p, \tau - \sigma)\tilde{V}_{\text{open}}(p, \tau + \sigma)$, where (τ, σ) are the world-sheet coordinates. This, in turn, implies that the tree-level n-closed-string amplitudes are given by sums of products,

$$\mathcal{A}^n_{\text{closed}} \sim \sum_{\sigma,\sigma'} e^{i\theta(\sigma,\sigma')}\mathcal{A}^n_{\text{open}}(\sigma)\tilde{\mathcal{A}}^n_{\text{open}}(\sigma') \tag{7}$$

where A^n_{open} and $\tilde{A}^n_{\text{open}}$ are left and right moving open string amplitudes, σ, σ' are noncyclic permutations of the external lines and the θ are model independent phases determined entirely by σ, σ' [34]. For example, focussing on the bosonic string consider the very simplest case of the three-graviton vertex,

$$\mathcal{A}^3_{\text{closed}} \sim \kappa\varepsilon_1^{\mu\alpha}\varepsilon_2^{\nu\beta}\varepsilon_3^{\rho\gamma} A^3_{\text{open }\mu\nu\rho}\tilde{A}^3_{\text{open }\alpha\beta\gamma}. \tag{8}$$

Here, κ is the closed string coupling constant, $\varepsilon_i^{\mu\alpha}$ are the transverse-traceless graviton polarisation tensors and the open string three-gluon vertex is given by

$$\mathcal{A}^3_{\text{open}} \sim g\varepsilon_1^{\mu}\varepsilon_2^{\nu}\varepsilon_3^{\rho} A^3_{\text{open }\mu\nu\rho}, \tag{9}$$

where g is the open string coupling constant and ε_1^{μ} are the transverse gluon polarisation tensors. Note, at zeroth-order in the inverse string tension, α', $\mathcal{A}^3_{\text{open}}$ is precisely the three-gluon vertex of pure Yang–Mills theory, but also has an order α' contribution corresponding to an F^3 term. Consequently, $\mathcal{A}^3_{\text{closed}}$ has both order α' and α'^2 contributions following from four- and six-derivative terms of the form R^2 and R^3. However, in the infinite string tension limit, $\alpha' \to 0$, we recover a precise relationship between the three-point vertices of perturbative $\mathcal{N} = 0$ supergravity (which can be restricted to Einstein–Hilbert gravity by choosing the gluon polarisations appropriately, since it is tree-level) and Yang–Mills theory. At four-point with external tachyons we obtain a relationship [34] between the famous closed string Virasoro–Shapiro [37,38] and open string Veneziano [39] amplitudes, which initiated the string theory programme itself. The four-point KLT relations hold for all external states and, again, reduce to four-point Einstein–Hilbert gravity and Yang–Mills relations in the $\alpha' \to 0$ limit. More generally, (7) for $\alpha' \to 0$ gives a precise set of relations between graviton amplitudes as sums of products of gluon amplitudes for any number of points [34].

The KLT relations and their field theory descendants are by construction tree-level. Nonetheless, the "closed = open × open" string theoretic approach was successfully applied to one-loop four-graviton amplitudes [40], suggesting that it may be possible to go beyond the semi-classical regime. At this stage it should be noted that direct contact with the standard Lagrangian approach has been lost. Instead, the key idea is to use unitarity methods [41–47] to build from the tree-level KLT relations loop amplitudes without passing through the usual Feynman prescription at all. This facilitated, for example, the calculation of the two-loop four-points amplitudes in $D = 4, \mathcal{N} = 8$ supergravity from those of $\mathcal{N} = 4$ super Yang–Mills theory [48]. The idea that unitarity can be used to glue trees into loops forms a part of what might be described as the "on-shell paradigm". Starting with Lagrangian field theory we learnt how to perturbatively compute amplitudes to arbitrary precision. While conceptually straightforward, the factorial growth in complexity with loop order quickly renders the traditional Feynman diagram approach impractical. The search for computational efficiency precipitated a renaissance in amplitude techniques, focusing on physical, gauge invariant objects. Over time various amplitude structures (recursion relations, generalised unitarity cuts, Grassmannians, scattering equations, positively…) were uncovered, eventually allowing the Lagrangian ladder to be kicked away altogether. As well as being computationally powerful, this programme offers new perspectives on quantum field theory itself. For a review of many of these developments see [49].

Importantly, this new found freedom led to the discovery of new features of amplitudes, not visible from the original Lagrangian perspective. A remarkable example of one such hidden structure, that lies at the heart of "gravity $=$ gauge $\times$ gauge", is the Bern–Carrasco–Johansson colour-kinematic (BCJ) duality, which relates the kinematic dependence of an amplitude to its colour data [50,51]. One can write any gluon amplitude entirely in terms of trivalent graphs (which are *not* Feynman diagrams) by "blowing" up the four-point contact terms. Having done so, the BCJ duality conjecture is that there exists a rewriting of the amplitude such that: (i) for any triple of graphs, i, j, k with colour factors c_i, c_j, c_k, built entirely from the gauge group structure constants, obeying a Jacobi identity $c_i + c_j + c_k = 0$, the corresponding kinematic factors, n_i, n_j, n_k, which are built from the momenta and polarisation tensors, also obey the same Jacobi-type identity $n_i + n_j + n_k = 0$; (ii) for any diagram, i, such that $c_i \to -c_i$ under the interchange of two legs then $n_i \to -n_i$. A reorganisation admitting this surprising relationship between colour and kinematic data exists for all n-point tree-level amplitudes, as has been demonstrated from a number of perspectives [52–55]. Although there is as yet no proof that the colour-kinematic duality will hold to all loops, there are many highly non-trivial examples providing supportive evidence [56–63]. While it is clear that the colour factors should obey Jacobi identities (by definition), it is not at all obvious that the kinematic factors should play by the same rules! It is certainly not apparent form the Yang–Mills Lagrangian. Moreover, they have further implications for amplitude architecture. For example, an immediate consequence of tree-level BCJ duality is the existence of BCJ *relations* amongst colour-ordered partial amplitudes, reducing the number of independent n-point partial amplitudes down to $(n-3)!$ [50].

More remarkable still is the BCJ double-copy prescription [51,64]. Consider two n-point L-loop Yang–Mills amplitudes, both written in trivalent form with respective colour and kinematic factors (c_i, n_i) and $(c_i, \tilde{n}_i)$, at least one of which has been successfully cast in a BCJ duality respecting form, say (c_i, n_i). We can construct a corresponding gravitational theory amplitude by simply replacing each colour factor in $(c_i, \tilde{n}_i)$ with the corresponding kinematic factor of (c_i, n_i): $(c_i, \tilde{n}_i) \to (n_i, \tilde{n}_i)$. We have removed all reference to the gauge group and "doubled" the kinematic terms.

For two pure Yang–Mills theories[7] this double-copy procedure generates all possible amplitudes of $\mathcal{N} = 0$ supergravity, giving precise meaning to the heuristic equation (5), at least at the semi-classical level. However, the two amplitudes need not belong to the same theory. For example, we could take the (c_i, n_i) from maximally supersymmetric $\mathcal{N} = 4$ Yang–Mills amplitudes and the $(\tilde{c}_i, \tilde{n}_i)$ from pure Yang–Mills theory. This yields the amplitudes of pure $\mathcal{N} = 4$ supergravity [65]. Alternatively, if both factors are $\mathcal{N} = 4$ Yang–Mills theory we generate the maximally supersymmetric $\mathcal{N} = 8$ supergravity [51], which can be thought of as the dimensional reduction on a 6-torus of the "type II = type I × type I" relation described in (6). By varying the left and right factors over all BCJ duality compatible gauge theories we generate all BCJ double-copy constructible gravitational theories. Of course, this is easier said than done, but there is nonetheless a rapidly multiplying zoology BCJ double-copy constructible gravity theories [51,56,64,66–89]. We shall come to describe this forest of theories once we have covered the necessary ground work.

The double-copy picture is not only conceptually compelling but also computationally powerful, bringing previously intractable calculations with in reach. This has pushed forward dramatically our understanding of divergences in perturbative quantum gravity [56,57,60,67,90–95], revealing a number of unexpected features and calling into question previously accepted arguments regarding finiteness. A remarkable example is given by the four-point graviton amplitude in $\mathcal{N} = 8$ supergravity, which was shown to be finite to four loops in [56], contradicting some early expectations [96,97]. It has since been shown that the four-loop cancellation can be accounted for by supersymmetry and $E_{7(7)}$ U-duality [98–102]. The consensus, however, is that at seven loops any would-be cancellations cannot be "consequences of supersymmetry in any conventional sense" [98]. Unfortunately, seven loops in $\mathcal{N} = 8$ supergravity remains beyond reach (for now), but by decreasing the amount of supersymmetry the same arguments apply at lower loop order. For example, the four-point amplitude of $D = 4, \mathcal{N} = 5$ supergravity has been shown to be finite to four loops, contrary to all expectations based on standard symmetry arguments [57]. There are "enhanced cancellations" [57], in the sense that they cannot be explained by any standard symmetry argument[8], at work and the conclusion that $\mathcal{N} = 8$ supergravity will diverge at seven loops is thrown into doubt. More recently, in a computational *tour de force* the $\mathcal{N} = 8$ four-point five-loop amplitude was completed using generalised BCJ duality and the double-copy of $\mathcal{N} = 4$ Yang–Mills theory [95,104]. It was found to be finite in agreement with the expectation that $\mathcal{N} = 8$ should diverge at seven loops. Its degree of finiteness was inline with standard symmetry arguments; there were no enhanced cancellations that might lead one to speculate that seven loops will be finite contrary to conventional expectations. However, this conclusion was reached in $D = 24/5$ where the five-loop amplitude first diverges and there is a $\partial^8 R^4$ counter-term. Since we do not fully understand the nature or origin of the enhanced cancellations it could still be that they kick in at seven loops in the relevant of $D = 4$. Similarly, the fact that $\mathcal{N} = 5$ supergravity is rendered finite at four loops by enhanced cancellations

[7] With possibly distinct gauge groups. The construction did not rely on any particular properties of, or relations amongst, the gauge groups.

[8] See [103] for possible explanations at three loops that nonetheless fail at four loops.

can be interpreted in a number of ways. On the one hand, it makes it clear that standard symmetry arguments are insufficient to predict when a supergravity theory will diverge, opening a small window of opportunity for perturbative finiteness. On the other hand, although not a logical impossibility no amplitude practitioners (as far as we are aware) expect $\mathcal{N} = 5$ supergravity to be perturbatively finite so: (i) the $\mathcal{N} = 5$ enhanced cancellations witnessed at four loops are anticipated to fail eventually and even if $\mathcal{N} = 8$ supergravity (where there *is* active debate regarding finiteness, see for example [98,105–108]) is magically finite at seven loops, it is not guaranteed that the cancellations will persist at higher loops; (ii) the observation of enhanced cancellations in $\mathcal{N} < 8$ supergravity theories implies that $\mathcal{N} = 8$ is not actually special in this regard. Does this undermine its privileged position as a candidate finite theory; if theories that are not expected to be finite can have enhanced cancellations, then why should we think that their existence might suggest finiteness of other theories? Of course, without a complete understanding of the amplitudes, including hidden features such as the enhanced cancellations, these are all just speculations and we will not know the answer at any particular order until we do the calculation.

Whatever the case, however, there is something deeper at work we have yet to fully comprehend; the questions regarding finiteness and the "gravity = gauge × gauge" paradigm more generally, remain very much open. In particular:

1. Why does the correspondence work? Can we prove the BCJ colour-kinematic conjecture? Is there some underlying geometric or world-sheet origin?
2. How deep is the correspondence? Is "gravity = gauge × gauge" strictly a property of amplitudes or can it be generalised to other/all aspects of gauge and gravity theories? What are the implications for quantum gravity?
3. How general is the correspondence? When does a gauge theory respect BCJ duality? What classes of gravitational theories admit a gauge theory squared origin; are the factorisable theories special in some regard?

1.1 Outline

In the remainder of this review we shall begin by describing BCJ duality and the double-copy construction in some detail before giving an (inevitably incomplete) overview of the rapidly evolving work tackling these questions and their various related puzzles. In Sect. 2 we consider "gravity = gauge × gauge" in the context of scattering amplitudes, where it takes its most concrete and developed form. In particular, we shall use this section to introduce some notation and the basic background concepts, before introducing BCJ duality and the double-copy construction in some detail. This will address to some degree question (3) above regarding what theories admit a "gravity = gauge × gauge" origin. We shall also review the status of the implications for perturbative quantum gravity.

Finally in Sect. 3, we take a step back from amplitudes and discuss various "off-shell" approaches to understanding BCJ duality and "gravity = gauge × gauge", addressing aspects of (1), (2) and (3). The first example is a Lagrangian approach to making BCJ duality manifest introduced in the original work on the double-copy [51,64]. One can also consider the double copy of classical gauge solutions [109–

135]. We will then review a field theoretic formulation of "gravity $=$ gauge $\times$ gauge" independent, but consistent with, the BCJ double-copy [81,85,115,116,119,136–142].

A comment on the scope of this review. Given the length constraints of La Rivista we have not been able to be pedagogical nor close to comprehensive. Regarding the latter, we can only apologise for any omissions and would welcome suggestions for future additions. Regarding the former, while it has not been possible to be pedagogical, we have endeavoured to be reasonably self-contained and elementary on the essential introductory matter so the non-expert will have some chance to understand the core basics, opening the door as it were. Otherwise, further reading will no doubt be required and we have tried to provide sufficient references throughout.

Before proceeding any further let us first mention other reviews that cover the various subthemes treated here in more detail. First, for a pedagogical introduction to BCJ duality and the double-copy one could not do better than [143,144]. The latter is also the superlative reference for BCJ duality, the double-copy and their applications more generally. For students starting out in this subject, or those coming from outside the amplitudes community, [44,49,145–148] provide excellent introductions, the latter two with one eye on BCJ duality the double-copy from the outset. For a broader, rather inspirational, early account of "gravity $=$ gauge $\times$ gauge" and its potential implications for perturbative quantum gravity see [149]. For an eminently approachable review of twistors, setting the scene for their applications to amplitudes, see [150]. For an excellent account of their applications to perturbative quantum field theory and the relationship between gauge and gravity amplitudes see [151].

2 Scattering amplitudes

The slogan "gravity $=$ gauge $\times$ gauge" takes its most concrete and complete form in the setting of scattering amplitudes and so this is where we shall begin our journey. In the context of modern particle physics tested at accelerators, such as the Large Hadron Collider, scattering amplitudes constitute the most basic gauge theory observables. Through their relation to cross-sections for scattering processes, they encode the probabilities that a set of colliding particles will interact to produce some other set of particles, thus providing direct contact between theory and experiment. Belying their conceptual simplicity, they are replete with subtle hidden structures and continue to reveal remarkable surprises to this day, with profound consequences, not only for computational techniques, but also quantum field theory itself.

One such surprise is the idea that gauge theory amplitudes can be used as the building blocks for gravity theory amplitudes. This approach can be traced back to the KLT relations of string theory. The modern, further reaching, incarnation takes the form of the BCJ double-copy construction, which will be our principal preoccupation here. The double-copy in-turn relies upon BCJ colour-kinematic duality, which is thus our first concern.

2.1 Bern–Carrasco–Johansson colour-kinematic duality

The cornerstone of the double-copy realisation of graviton scattering amplitudes in terms of gluon scattering amplitudes is the Bern–Carrasco–Johansson *colour-kinematic duality* [50]. Everyone knows that the colour factors dressing gluonic Feynman diagrams obey certain relations, such as the Jacobi identity. The colour factors derive from the Lie group characterising the gauge theory so it is in their very nature to do so. However, BCJ duality implies that there exists a rewriting of the amplitude such that whenever a set of diagrams satisfies a Jacobi identity amongst their colour factors, the corresponding kinematic numerators obey precisely the same identity. Kinematic numerators satisfying these identities are referred to BCJ numerators. It is not at all obvious that this should be true; it certainly is not manifest in the conventional Yang–Mills Lagrangian.

Let us now describe in detail the BCJ duality conjecture for pure Yang–Mills theory. We shall add supersymmetry, matter couplings and more exotic structures in the subsequent sections. For now, we take this opportunity to give a lightning review of Yang–Mills theory, setting some notation, and introducing the basic ideas that will carry us through the remaining material.

2.1.1 Yang–Mills theory and gluon scattering amplitudes

We begin by reviewing pure Yang–Mills gauge theories, which are specified by a choice of compact Lie group G and a principal G-bundle $P(M, G)$, defined over a base manifold M corresponding, in this context, to a fixed spacetime background. Given an open subset U of M with local section σ, the local Yang–Mills gauge potential $A \in \Omega^1(U) \otimes \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of G, is given by $A = \sigma^*\omega$, where ω is a connection on $P(M, G)$. The associated Yang–Mills field strength given by

$$F = dA + A \wedge A \tag{10}$$

then corresponds to a local form of the curvature of the connection. For an open covering $\{U_i\}$ of M with local sections σ_i, the gauge potentials on non-trivial overlaps $U_i \cap U_j \neq \varnothing$ satisfy the compatibly relations,

$$A_j = t_{ij}^{-1} A_i t_{ij} + t_{ij}^{-1} dt_{ij}, \tag{11}$$

where $t_{ij} : U_i \cap U_j \to G$ are the transition functions of $P(M, G)$. This implies the compatibility relation for the associated field strengths

$$F_j = t_{ij}^{-1} F_i t_{ij}. \tag{12}$$

For two local sections on a chart U related by $\sigma'(p) = \sigma(p)g(p)$ for all $p \in U$, where $g : U \to G$, the corresponding gauge potentials are related by the gauge transformations

$$A' = g^{-1} A g + g^{-1} dg, \tag{13}$$

which imply

$$F' = g^{-1} F g. \tag{14}$$

For gauge transformations connected to the identity there is a $\theta : U \to \mathfrak{g}$ such that $g = \exp[\theta]$ and

$$\delta A = d\theta + [A, \theta] \equiv D_A \theta, \quad \delta F = [F, \theta] \tag{15}$$

to first order in θ, where we define the commutator[9] of $\mathfrak{g}$-valued forms by $[x, y] = x \wedge y - (-1)^{pq} y \wedge x$, for $x \in \Omega^p(U) \otimes \mathfrak{g}$ and $y \in \Omega^q(U) \otimes \mathfrak{g}$.

Here we have introduced the local covariant derivative $D_A : \Omega^p(U) \otimes \mathfrak{g} \to \Omega^{p+1}(U) \otimes \mathfrak{g}$ for local connection A (we will henceforth omit the subscript on D_A indicating the choice of connection),

$$Dx := dx + [A, x]. \tag{16}$$

The field strength obeys the local Bianchi identity,

$$DF = d(dA + A \wedge A) + A \wedge dA - dA \wedge A = 0. \tag{17}$$

It is often useful to introduce a basis $\{T_a\}_{a=1}^{\dim \mathfrak{g}}$ for $\mathfrak{g}$ in some representation ρ, especially when considering the colour structure of scattering amplitudes. Then

$$A = g A^a_\mu dx^\mu \otimes T_a, \quad F = \frac{1}{2} g F^a_{\mu\nu} dx^\mu \wedge dx^\nu \otimes T_a \tag{18}$$

where g is the Yang–Mills coupling and

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a{}_{bc} A^b_\mu A^c_\nu \tag{19}$$

and $f^a{}_{bc}$ are the structure constants of $\mathfrak{g}$, $[T_a, T_b] = f_{ab}{}^c T_c$, which are totally antisymmetric $f_{abc} = f_{[abc]}$ (adjoint indices are raised/lowered via minus the Cartan-Killing form, so may be taken as δ_{ab} in an appropriate basis) and satisfy the Jacobi identity

$$3 f_{[ab}{}^e f_{c]e}{}^d = f_{ab}{}^e f_{ce}{}^d + f_{ca}{}^e f_{be}{}^d + f_{bc}{}^e f_{ae}{}^d = 0. \tag{20}$$

In components the gauge transformations (15) are given by

$$\delta A^a_\mu = \partial_\mu \theta^a + g f^a{}_{bc} A^b_\mu \theta^c, \quad \delta F^a_{\mu\nu} = g f^a{}_{bc} F^b_{\mu\nu} \theta^c. \tag{21}$$

where $\theta = g\theta^a T_a$. With the goal of developing gluonic scattering amplitude relations in mind, for the remainder of this section we shall restrict M to be $D = (1 + d)$ dimensional Minkowski spacetime so that the bundle is trivial. We may leave for now G to be an arbitrary compact Lie group, since the adjoint representation is always real, but the typical example to have in mind is SU(N). When we consider matter couplings we will have to be more careful about the properties of G, taking into

[9] This is only unambiguous for matrix Lie algebras, but the general case is given by the obvious interpretation for an arbitrary Lie bracket $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$.

account the specific reality properties of the other representations required. With these comments in mind, we can now turn to the perturbative quantum Yang–Mills theory valid in the high energy regime.

The quantum theory is most transparently formulated through an action principle. The classical Yang–Mills action functional is given by

$$S_{\rm YM} = \frac{1}{2g^2}\int_M \operatorname{tr} F \wedge \star F = -\frac{1}{4}\int_M d^D x\, F^a_{\mu\nu} F^{a\mu\nu} \tag{22}$$

where tr denotes an appropriately normalised G-invariant and negative-definite quadratic form on $\mathfrak{g}$. It is by construction invariant under the gauge transformations (15). Consequently, there is a large gauge redundancy, which must be treated carefully using, for example, the Faddeev–Popov proceedure [152]. Although well-trodden territory, we will briefly review the approach of Becchi–Rouet–Stora–Tyutin (BRST) quantisation [153,154], as certain ingredients will be explicitly needed later in the perhaps less familiar context of a *classical* field-theoretic realisation of "gravity = gauge × gauge". There are several good reviews, for example [155–157], of the BRST formalism and the more general Batalin-Vilkovisky (BV) [158–163] approach, as needed for open symmetry algebras encountered, for example, in supergravity. We will in fact need the full machinery of the BV formalism later, but refer the reader to [157] for the required background material.

Following the Faddeev–Popov prescription, using the Nakanishi–Lautrup Lagrange multiplier $b : M \to \mathfrak{g}$ we can lift to the action the generalised gauge-fixing delta $\delta(G[A] - w)$, with a width ξ Gaussian weighting by $w : M \to \mathfrak{g}$. The Faddeev–Popov determinant is also lifted to the action by the inclusion of the anti-commuting ghost and antighost fields $c, \bar{c} : \Omega^0(M) \to \mathfrak{g}$. This results in the total Yang–Mills BRST action

$$S_{\rm YM_{BRST}} = S_{\rm YM} + S_{\rm gf} + S_{\rm gh} \tag{23}$$

where

$$S_{\rm gf} = -\frac{1}{g^2}\int_M \operatorname{tr} b\left(\frac{1}{2}\xi b + G[A]\right) \tag{24}$$

follows from the gauge-fixing terms and

$$S_{\rm gh} = -\frac{1}{g^2}\int_M d^4 x \operatorname{tr}\left(\bar{c}\int_M d^4 y\, \tfrac{\delta G(x)}{\delta A^\mu(y)} D^\mu c(y)\right), \tag{25}$$

follows from the Faddeev–Popov determinant. Although we have explicitly broken the gauge symmetry by the addition of (24), the total Yang–Mills BRST action (23) is annihilated under the global BRST transformations, in which the ghost field replaces the local gauge parameter θ:

$$QA := Dc, \tag{26a}$$

$$Qc := -\tfrac{1}{2}\{c, c\}, \tag{26b}$$

$$Q\bar{c} := b, \tag{26c}$$

$$Qb := 0, \tag{26d}$$

where $Q, c, \bar{c}, b$ have ghost numbers $\mathrm{gh}(Q) = 1, \mathrm{gh}(c) = 1, \mathrm{gh}(\bar{c}) = -1, \mathrm{gh}(b) = 0$ and

$$Q(ab) = (Qa)b + (-1)^{\varepsilon(a)} a(Qb) \tag{27}$$

where $\varepsilon(f) \in \{0, 1\}$ is the Grassmann grade of f.

In this case the BRST charge is nilpotent, $Q^2 = 0$, without imposing any further conditions on the fields, which follows from the off-shell closure of the classical gauge transformations. It is the Noether charge associated to the continuous global symmetry with anti-commuting parameter ϵ, given by $\delta_\epsilon = \epsilon Q$. Using these transformations we can rewrite $S_{\mathrm{gf}} + S_{\mathrm{gh}}$ in terms of a Q-exact term $\int \star Q \Psi_{\mathrm{gf}}$, where the ghost-fixing fermion Ψ_{gf} is given by $-\mathrm{tr}\, \bar{c}(\frac{1}{2}\xi b + G[A])$. Functions on this enlarged space of fields, $(A, c, \bar{c}, b)$, together with the homological vector field, Q, form a chain complex and its cohomology characterises the set of possible physical observables. The expectation values of physical observables are independent of the choice of gauge-fixing, which is made manifest by the freedom to add Q-exact terms that modify the gauge-fixing and ghost terms, but leave the physical sector invariant.

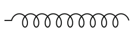
Making a choice for $G[A]$ and eliminating the auxiliary field b, we can then follow the usual perturbative approach to quantising (23), safe in the knowledge that the gauge redundancy has been appropriately accounted for, to arrive at the familiar Feynman diagrams for pure Yang–Mills theory, as described in the standard quantum field theory textbooks.[10] Since the colour structure of these diagrams will lie at the heart of our discussion, let us be completely explicit nonetheless.

Consider the linear covariant gauge fixing condition $G[A] = \mathrm{div} A$. After eliminating the Nakanishi-Lautrup field b through its algebraic equation of motion we are left with,

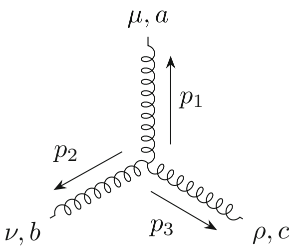
$$S_{\mathrm{YM_{BRST}}} = \frac{1}{2g^2} \int_M \mathrm{tr} \left(F \wedge \star F - \frac{1}{\xi} \star (d^\dagger A)^2 + 2 d\bar{c} \wedge \star Dc \right), \tag{28}$$

which has no gauge symmetry left but is invariant under the BRST transformations (26), but with $Q\bar{c} := -\frac{1}{\xi} d^\dagger A$, which follows from the equation of motion $b = -\frac{1}{\xi} d^\dagger A$.

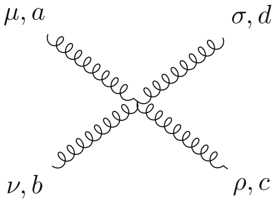
The Feynman diagrams then follow immediately from the standard canonical approach. The pure gluon diagrams in Feynman gauge ($\xi = 1$) are as follows:

ν, b  $\mu, a \qquad = -\dfrac{i\eta_{\mu\nu}\delta^{ab}}{p^2}$ (29a)

[10] An introduction particularly well-suited to the modern treatment of amplitudes is [164]. Coupled with [49] this will bring a graduate student up to speed for most purposes here.



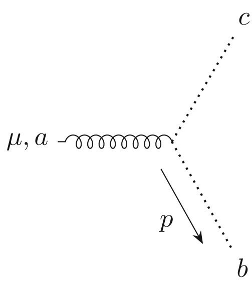
$$= g f^{abc}\left[p_{23}^{\mu}\eta^{\nu\rho} + p_{31}^{\nu}\eta^{\rho\mu} + p_{12}^{\rho}\eta^{\mu\nu}\right] \tag{29b}$$



$$= \begin{aligned} -ig^2\Big[& f^{abx}f_x{}^{cd}\eta^{\mu\nu\rho\sigma} \\ & + f^{adx}f_x{}^{bc}\eta^{\mu\sigma\nu\rho} \\ & + f^{acx}f_x{}^{db}\eta^{\mu\rho\nu\sigma}\Big] \end{aligned} \tag{29c}$$

where $p_{ij} = p_i - p_j$ and $\eta^{\mu\nu\rho\sigma} = \eta^{\mu\rho}\eta^{\nu\sigma} - \eta^{\mu\sigma}\eta^{\nu\rho}$. The ghost diagrams are given by,

$$= \frac{i\delta^{ab}}{p^2} \tag{30a}$$



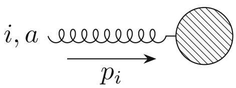
$$= -g f^{abc} p^{\mu} \tag{30b}$$

With the gluon Feynman diagrams at our disposal we can now turn to scattering amplitudes in pure Yang–Mills gauge theory. Consider a collection of n_i non-interacting and well-separated gluons in-coming from past infinity in the initial separable state

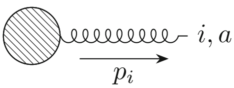
$$|\text{in}\rangle = |p_1, \varepsilon_1\rangle \otimes |p_2, \varepsilon_2\rangle \cdots \otimes |p_{n_i}, \varepsilon_{n_i}\rangle \equiv |p_1, \varepsilon_1; p_2, \varepsilon_2; \ldots p_{n_i}, \varepsilon_{n_i}\rangle, \tag{31}$$

where p_i, ε_i are the on-mass-shell momentum and polarisation (we are suppressing here the colour data) of the i^{th} gluon, respectively. The amplitude for these in-coming particles to scatter into some out-going collection of n_f non-interacting, well-separated particles at future infinity, $t \to \infty$, $|\text{out}\rangle = |p_{n_i+1}, \varepsilon_{n_i+1}; \ldots p_{n_i+n_f}, \varepsilon_{n_i+n_f}\rangle$, is given by the S-matrix element, $\mathcal{A}_{n_i \to n_f} = \langle \text{out}|S|\text{in}\rangle$.

The Lehman–Symanzik–Zimmermann reduction formula re-expresses the scattering amplitude $\mathcal{A}_{n_i \to n_f}$ in terms of the Green function of $n_i + n_f$ local fields dressed by polarisation tensors and reciprocal renormalised two-point Green functions, which precisely cancel the external propagators. Splitting the S-matrix into its trivial and interacting pieces $S = \mathbb{1} + iT$, the non-trivial amplitude is then given by the sum of all amputated and connected graphs using the Feynman diagrams in (29) contracted with the external legs, given in (32), with on-shell momenta $\{p_i\}_{i=1}^{n_i+n_f}$ and polarisations $\{\varepsilon_i\}_{i=1}^{n_i+n_f}$ (suppressing the helicity index $\varepsilon_i = \varepsilon_i^s$, $s = \pm$),



$$= \varepsilon_i^{\mu *}(p)\delta^{ab}, \qquad \varepsilon_i^* \cdot p_i = 0,\ p^2 = 0, \tag{32a}$$



$$= \varepsilon_i^{\mu}(p)\delta^{ab}, \qquad \varepsilon_i \cdot p_i = 0,\ p^2 = 0, \tag{32b}$$

where the μ, b indices are contracted with the colour and momentum indices of the associated leg of the internal diagram (represented here by the blob).

2.1.2 The colour-kinematic duality

Here we introduce the BCJ colour-kinematic duality for gluons. Let us first state the claim, so that we know where we are heading, before carefully unpacking the various ingredients.

First, recall an n-point L-loop gluon amplitude $\mathcal{A}_{\text{YM}}^{n,L}$ may always be written as

$$\mathcal{A}_{\text{YM}}^{n,L} = i^L g^{n-2+2L} \sum_i \int \prod_{l=1}^{L} \frac{d^D p_l}{(2\pi)^D S_i} \frac{c_i n_i}{d_i}, \tag{33}$$

where the sum is over all n-point L-loop graphs, labelled i, with only trivalent vertices (these are *not* Feynman diagrams). The *colour numerator* or *factor* c_i associated to graph i is composed of gauge group structure constants and can be read off directly from the graph. The *kinematic numerator* or *factor* n_i associated to graph i is a polynomial of Lorentz-invariant contractions of polarisation vectors and momenta. The denominators d_i are composed of the propagators given by products of the momentum squared of each internal line of graph i. Here, $S_i \in \mathbb{N}$ accounts for any over-counting due to the graph symmetries. At tree-level, $L = 0$, this simplifies to

$$\mathcal{A}^{n,0}_{\text{YM}} = g^{n-2} \sum_{i=1}^{(2n-5)!!} \frac{c_i n_i}{d_i}, \tag{34}$$

since there are $(2n-5)!!$ trivalent tree diagrams at n-points.[11]

Having reorganised the amplitudes into a sum over purely trivalent graphs we can state the BCJ duality conjecture:

BCJ colour-kinematic duality conjecture: there exists a choice of kinematic numerators of the trivalent diagrams entering the amplitude $\mathcal{A}^{n,L}_{\text{YM}}$ such that:

1. Whenever a triple of trivalent diagrams (i, j, k) has colour factors obeying the Jacobi identity

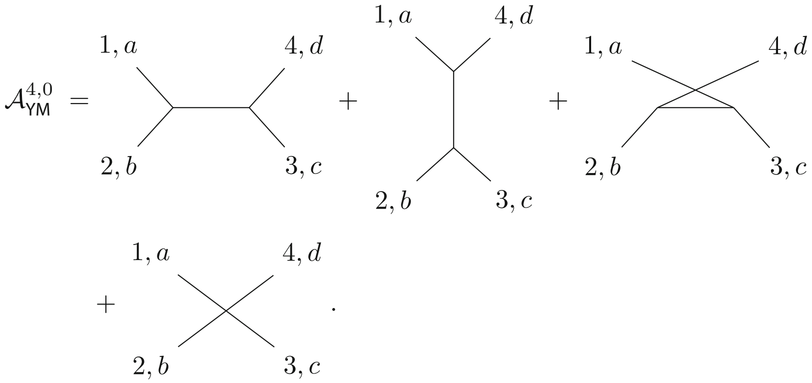
$$c_i + c_j + c_k = 0$$

then the corresponding kinematic factors obey precisely the same identity

$$n_i + n_j + n_k = 0.$$

2. If any individual diagram has $c_i \rightarrow -c_i$ under the interchange of two legs then $n_i \rightarrow -n_i$ at the same time.

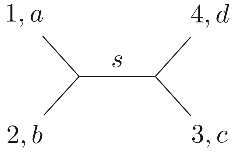
Having presented the conjecture, let us now expand on the various components, starting with (33). Since $[-,-] : \mathfrak{g} \otimes \mathfrak{g} \rightarrow \mathfrak{g}$, the possibility of relating colour to kinematics relies on writing the amplitude in terms of trivalent diagrams only. This is possible because the four-point contact terms (29c) can always be 'blown-up' and absorbed into three-point diagrams. Consider the simplest example of the four-point, tree-level amplitude,



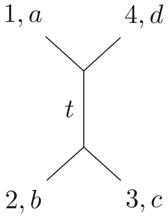
(35)

[11] Proceed by induction: clearly true for $n = 3$. At n-points each graph has $2n - 3$ edges to which one can add an edge to give $2n - 3$ new $(n + 1)$-point graphs. Hence, assuming there are $(2m - 5)!!$ graphs at n-points, there are $(2n - 3)(2n - 5)!! = (2(n + 1) - 5)!!$ graphs at $(n + 1)$-points.

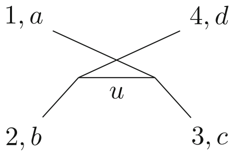
We shall leave the helicities unspecified as the argument should not depend on any particular configuration and simply denote the kinematic numerators[12] by $n = n(p_i, \varepsilon_i)$ as in (33). Explicitly, with all momenta outgoing $p_1 + p_2 + p_3 + p_4 = 0$,



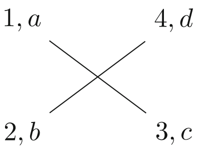
$$= -ig^2 \frac{f^{abx} f_x{}^{cd} n_s}{s} = -ig^2 \frac{c_s n_s}{s} \tag{36a}$$



$$= -ig^2 \frac{f^{axd} f_x{}^{bc} n_t}{t} = -ig^2 \frac{c_t n_t}{t} \tag{36b}$$



$$= -ig^2 \frac{f^{axc} f_x{}^{db} n_u}{u} = -ig^2 \frac{c_u n_u}{u} \tag{36c}$$



$$= -ig^2 \left[f^{abx} f_x{}^{cd} n_s^{(4)} + f^{adx} f_x{}^{bc} n_t^{(4)} + f^{acx} f_x{}^{db} n_u^{(4)} \right] \tag{36d}$$

where we have used the Mandelstam variables $s = (p_1 + p_2)^2$, $t = (p_1 + p_4)^2$, $u = (p_1 + p_3)^2$ in the trivalent s-, t- and u-channel diagrams, respectively. We have suggestively labelled the kinematic factors appearing in the four-point contact term by

[12] As an example, for the s-channel we have

$$n_s = 4\big(\varepsilon_1 \cdot p_2 \varepsilon_{2\lambda} - \varepsilon_2 \cdot p_1 \varepsilon_{1\lambda} + \tfrac{1}{2}\varepsilon_1 \cdot \varepsilon_2 p_{12\lambda}\big)\big(\varepsilon_3 \cdot p_4 \varepsilon_4^\lambda - \varepsilon_4 \cdot p_3 \varepsilon_1^\lambda + \tfrac{1}{2}\varepsilon_3 \cdot \varepsilon_4 p_{34}^\lambda\big).$$

Of course, fixing the helicities and making some sensible choices for the polarisations, this can be significantly simplified.

$n_s^{(4)}, n_t^{(4)}, n_u^{(4)}$, where explicitly

$$\begin{aligned} n_s^{(4)} &= 2\varepsilon_1 \cdot \varepsilon_{[3}\varepsilon_{4]} \cdot \varepsilon_2, \\ n_t^{(4)} &= 2\varepsilon_1 \cdot \varepsilon_{[2}\varepsilon_{3]} \cdot \varepsilon_4, \\ n_u^{(4)} &= 2\varepsilon_1 \cdot \varepsilon_{[4}\varepsilon_{2]} \cdot \varepsilon_3. \end{aligned} \tag{37}$$

Writing the colour factors corresponding to each trivalent diagram as

$$c_s = f^{abx} f_x{}^{cd}, \quad c_t = f^{axd} f_x{}^{bc}, \quad c_u = f^{axc} f_x{}^{db}, \tag{38}$$

the four-point contact term becomes

$$-ig^2 \left(c_s n_s^{(4)} - c_t n_t^{(4)} - c_u n_u^{(4)} \right) = -ig^2 \left(\frac{c_s s n_s^{(4)}}{s} - \frac{c_t t n_t^{(4)}}{t} - \frac{c_u u n_u^{(4)}}{u} \right) \tag{39}$$

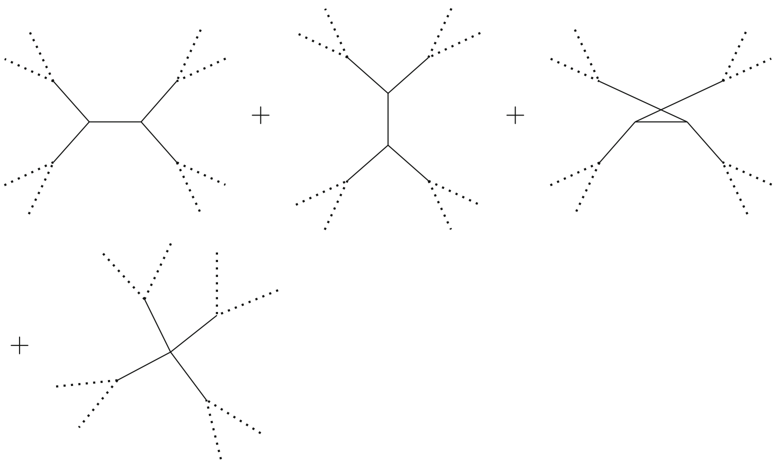
where we have trivially inserted the corresponding propagators. This makes it immediately clear that the three terms can be absorbed into the s-, t- and u-channels respectively, shifting their kinematic factors,

$$n_s \to n'_s = n_s + s n_s^{(4)}, \quad n_t \to n'_t = n_t - t n_t^{(4)}, \quad n_u \to n'_u = n_u - u n_u^{(4)}. \tag{40}$$

so that the amplitude is a sum over the three trivalent diagrams,

$$\mathcal{A}_{\text{YM}}^{4,0} = -ig_{\text{YM}}^2 \left(\frac{c_s n'_s}{s} + \frac{c_t n'_t}{t} + \frac{c_u n'_u}{u} \right). \tag{41}$$

This argument trivially goes through for any set of four arbitrarily complex diagrams that differ only by a set of four-point subdiagrams embedded in a common subsector as depicted here,



(42)

where the dotted lines connect each subdiagram to otherwise identical total diagrams. Hence, wherever we see a four-point contact term we can absorb it into the three corresponding diagrams with the trivalent s-, t- and u-channels in its place.

We have rather laboured this essentially trivial observation, that the three sets of structure constants appearing in the four-point contact term are the same as those in the three four-point trivalent diagrams, because it brings into focus the first ingredients of the BCJ duality. From the outset, it was clear that the kinematic numerators entering the Feynman diagram decomposition of the amplitude are not unique since the polarisations are only defined up to shifts, $\varepsilon(p) \to \varepsilon(p) + \alpha p$, which change each kinematic factor (but of course leave the amplitude itself invariant). This is no surprise, since each individual digram is not gauge invariant. However, the preceding discussion makes a second, less trivial, ambiguity in the kinematic numerators apparent, the *generalised gauge transformations* introduced [50]. The nomenclature derives from the observation that while they may look and feel like gauge transformations, there need not be any gauge transformation that actually realises a given generalised gauge transformation. To describe the generalised gauge transformations, let us return to our four-point example. Note that the three colour factors c_s, c_t, c_u of (36) are precisely the combinations of structure constants appearing in the Jacobi identity (20),

$$c_s - c_t - c_u = 3f^{xa[b}f_x{}^{cd]} = 0. \tag{43}$$

Hence, under a shift of the kinematic numerators by an arbitrary function α,

$$n_s \mapsto n_s - s\alpha, \quad n_t \mapsto n_t + t\alpha, \quad n_u \mapsto n_u + u\alpha, \tag{44}$$

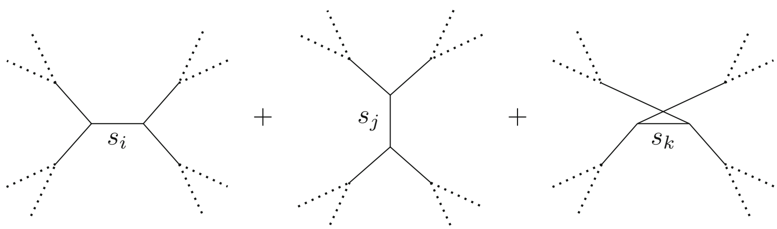
the amplitude (41) is left invariant,

$$\mathcal{A}^{4,0}_{\text{YM}} \mapsto -ig^2\left(\frac{c_s(n_s - s\alpha)}{s} + \frac{c_t(n_t + t\alpha)}{t} + \frac{c_u(n_u + u\alpha)}{u}\right) = \mathcal{A}^{4,0}_{\text{YM}}, \tag{45}$$

since $(-c_s + c_t + c_u)\alpha = 0$ by the Jacobi identity. Again, it is clear that this invariance generalises to any triple of trivalent diagrams (i, j, k) that only differ in a common four-point subsector with colour factor satisfying a Jacobi identity of the form $c_i + c_j + c_k = 0$, where the generalised gauge transformation acting on the corresponding kinematic numerators is given by,

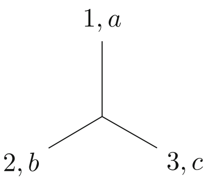
$$n_i \mapsto n_i + s_i\alpha, \quad n_j \mapsto n_j + s_j\alpha, \quad n_{|} \to n_k + s_k\alpha \tag{46}$$

and s_i, s_j, s_k are the three (and only three) distinct propagators as illustrated here,



Let us summarise. Every gluon scattering amplitude can be written in terms of purely trivalent graphs. The kinematic numerators associated to these graphs are not unique. In particular, for any triple of such diagrams with colour factors obeying a Jacobi identity, the amplitude is invariant under the generalised gauge transformations acting on the corresponding kinematic numerators. The BCJ colour-kinematic conjecture states that there is a writing of the kinematic numerators, exploiting their ambiguity, such that (i) whenever the colour factors of a triple of graphs obey a Jacobi identity then so do the corresponding kinematic numerators and, (ii) if interchanging two legs of diagram i implies $c_i \mapsto -c_i$, then we also have $n_i \mapsto -n_i$.[13] Before discussing the conjecture further let us take a look at some simple examples to get a feel for it.

We start with a triviality, the three-point amplitude (allowing complex momenta), which consists of a single diagram,



$$= gf^{abc}\left[\varepsilon_1 \cdot p_{23}\varepsilon_2 \cdot \varepsilon_3 + \varepsilon_2 \cdot p_{31}\varepsilon_3 \cdot \varepsilon_1 + \varepsilon_3 \cdot p_{12}\varepsilon_1 \cdot \varepsilon_2\right] = gcn. \tag{47}$$

Under interchange of any two edges $c = f^{abc} \mapsto -c$ since f^{abc} is totally antisymmetric. Since $p_{ij} = -p_{ji}$, we see that the same is true for n, as claimed.

The next example, tree-level four points, is already less immediately obvious, although it has been known to satisfy BCJ duality for some time now (before the notion of BCJ duality had been articulated) [165,166]. From (36), (37) and (40) the kinematic numerators with momenta out-going in the trivalent form are given by

$$\begin{aligned}
n_s &= 4\big(\varepsilon_1 \cdot p_2\varepsilon_2 - \varepsilon_2 \cdot p_1\varepsilon_1 + \tfrac{1}{2}\varepsilon_1 \cdot \varepsilon_2 p_{12}\big) \cdot \big(12 \to 34\big) + 2s\varepsilon_1 \cdot \varepsilon_{[3}\varepsilon_{4]} \cdot \varepsilon_2 \\
n_t &= 4\big(\varepsilon_4 \cdot p_1\varepsilon_1 - \varepsilon_1 \cdot p_4\varepsilon_4 + \tfrac{1}{2}\varepsilon_4 \cdot \varepsilon_1 p_{41}\big) \cdot \big(41 \to 23\big) + 2t\varepsilon_4 \cdot \varepsilon_{[2}\varepsilon_{3]} \cdot \varepsilon_1 \\
n_u &= 4\big(\varepsilon_4 \cdot p_2\varepsilon_2 - \varepsilon_4 \cdot p_4\varepsilon_4 + \tfrac{1}{2}\varepsilon_4 \cdot \varepsilon_2 p_{42}\big) \cdot \big(42 \to 31\big) + 2u\varepsilon_4 \cdot \varepsilon_{[3}\varepsilon_{1]} \cdot \varepsilon_2
\end{aligned} \tag{48}$$

Recall, the perhaps unfamiliar final terms come from absorbing the four-point contact term, hence the appearance of the propagators s, t, u. The claim is that since $c_s - c_t - c_u = 0$, see (43), by BCJ duality (without any further intervention in this case) we also have

$$n_s - n_t - n_u = 0 \tag{49}$$

on-shell ($\sum_{i=1}^4 p_i = 0$, $p_i^2 = 0$, $\varepsilon_i \cdot p_i = 0$). Using

$$\begin{aligned}
p_{12} \cdot p_{34} &= u - t, \\
p_{41} \cdot p_{23} &= u - s, \\
p_{42} \cdot p_{31} &= s - t,
\end{aligned} \tag{50}$$

[13] Note, the Kleiss–Kuijf relations [17,18] hold if the colour-ordered partial amplitudes are written as a sum of numerators satisfying this condition.

we see that the three unfamiliar terms, deriving from the four-point contact term, cancel identically against the terms $\varepsilon_1 \cdot \varepsilon_2 \varepsilon_1 \cdot \varepsilon_2 p_{12} \cdot p_{34}$, $\varepsilon_4 \cdot \varepsilon_1 \varepsilon_2 \cdot \varepsilon_3 p_{41} \cdot p_{23}$ and $\varepsilon_4 \cdot \varepsilon_2 \varepsilon_3 \cdot \varepsilon_1 p_{42} \cdot p_{31}$. To handle the remaining 24 terms we can first make a judicious choice for the polarisation reference vectors, $q_1 = q_2 = q_3 = p_4$ and $q_4 = p_2$, so that

$$\varepsilon_i \cdot p_4 = 0, \quad \varepsilon_4 \cdot p_2 = 0. \tag{51}$$

There are also various vanishing products amongst the polarisation tensors, for example $\varepsilon_i^{\pm} \cdot \varepsilon_j^{\pm} = 0, \forall i, j = 1, 2, 3$, but we shall not need these as the BCJ duality holds for all helicity configurations. By inspection we see that for this choice $n_u = 0$ and we are left with

$$\begin{aligned} n_s - n_t = 4\Big[&\varepsilon_4 \cdot p_3 \left(\varepsilon_2 \cdot p_1 \varepsilon_1 \cdot \varepsilon_3 - \varepsilon_1 \cdot p_2 \varepsilon_2 \cdot \varepsilon_3 - \tfrac{1}{2}\varepsilon_1 \cdot \varepsilon_2 \varepsilon_3 \cdot p_{12}\right) \\ &+ \tfrac{1}{2}\varepsilon_3 \cdot \varepsilon_4 (\varepsilon_1 \cdot p_2 \varepsilon_2 \cdot p_3 - \varepsilon_2 \cdot p_1 \varepsilon_1 \cdot p_3)\Big] \\ &- 4\Big[\varepsilon_4 \cdot p_1 \left(\varepsilon_2 \cdot p_3 \varepsilon_3 \cdot \varepsilon_1 - \varepsilon_3 \cdot p_2 \varepsilon_2 \cdot \varepsilon_1 - \tfrac{1}{2}\varepsilon_3 \cdot \varepsilon_2 \varepsilon_1 \cdot p_{32}\right) \\ &+ \tfrac{1}{2}\varepsilon_1 \cdot \varepsilon_4 (\varepsilon_3 \cdot p_2 \varepsilon_2 \cdot p_1 - \varepsilon_2 \cdot p_3 \varepsilon_3 \cdot p_1)\Big] \end{aligned} \tag{52}$$

It is straightforward to show that this combination vanishes due to the special four-point kinematics (with our choice of polarisation reference vectors, but of course there is no loss of generality). First, note that $\varepsilon_2 \cdot p_1 \varepsilon_1 \cdot p_3 = \varepsilon_1 \cdot p_2 \varepsilon_2 \cdot p_3$ since $\varepsilon_1 \cdot p_3 = \varepsilon_1 \cdot (p_3 + p_1) = -\varepsilon_1 \cdot (p_2 + p_4) = -\varepsilon_1 \cdot p_2$ and $\varepsilon_2 \cdot p_1 = \varepsilon_2 \cdot (p_1 + p_2) = -\varepsilon_2 \cdot (p_3 + p_4) = -\varepsilon_2 \cdot p_3$. Consequently, $\varepsilon_1 \cdot p_2 \varepsilon_2 \cdot p_3 - \varepsilon_2 \cdot p_1 \varepsilon_1 \cdot p_3 = 0$ and similarly $\varepsilon_3 \cdot p_2 \varepsilon_2 \cdot p_1 - \varepsilon_2 \cdot p_3 \varepsilon_3 \cdot p_1 = \varepsilon_4 \cdot p_3 \varepsilon_2 \cdot p_1 - \varepsilon_4 \cdot p_1 \varepsilon_2 \cdot p_3 = 0$, leaving

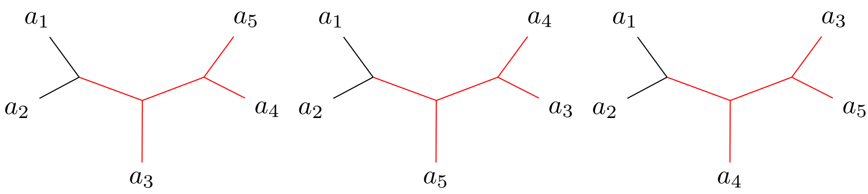
$$\begin{aligned} n_s - n_t &= 4\Big[\varepsilon_2 \cdot \varepsilon_1 (\varepsilon_4 \cdot p_1 \varepsilon_3 \cdot p_2 - \tfrac{1}{2}\varepsilon_4 \cdot p_3 \varepsilon_3 \cdot p_{12}) \\ &\qquad + \varepsilon_2 \cdot \varepsilon_3 (\varepsilon_1 \cdot p_2 \varepsilon_4 \cdot p_3 - \tfrac{1}{2}\varepsilon_4 \cdot p_1 \varepsilon_1 \cdot p_{32})\Big] \\ &= 0 \end{aligned} \tag{53}$$

where we have reorganised the remaining terms to make the final cancellations clear. Using $\varepsilon_4 \cdot p_3 = \varepsilon_4 \cdot (p_3 + p_4) = -\varepsilon_4 \cdot (p_1 + p_2) = -\varepsilon_4 \cdot p_1$, we have $\varepsilon_4 \cdot p_1 \varepsilon_3 \cdot p_2 - \frac{1}{2}\varepsilon_4 \cdot p_3 \varepsilon_3 \cdot p_{12} = \frac{1}{2}\varepsilon_4 \cdot p_1 \varepsilon_3 \cdot (p_1 + p_2) = -\frac{1}{2}\varepsilon_4 \cdot p_1 \varepsilon_3 \cdot (p_3 + p_4) = 0$ and similar $\varepsilon_1 \cdot p_2 \varepsilon_4 \cdot p_3 - \frac{1}{2}\varepsilon_4 \cdot p_1 \varepsilon_1 \cdot p_{32} = 0$. While this four-point example can be accounted for by the special kinematics associated to four-points, it makes the principle clear.

Five-points, considered the original BCJ duality work [50], provides the first truly non-trivial check. It also makes the general principles entering the duality clear, so let us reexamine it here, following closely [50], but of course with the benefit of hindsight. At five points there are 15 trivalent diagram contributing to the amplitude

$$\mathcal{A}_{\text{YM}}^{5,0} = \sum_{i=1}^{15} \frac{n_i c_i}{d_i} \tag{54}$$

Any triple of diagrams with a common pair of joined external legs gives us a colour Jacobi identity, for example



(55)

which corresponds to the identity

$$f^{a_1 a_2}{}_x \left(f^{x a_3 y} f_y{}^{a_4 a_5} + f^{x a_5 y} f_y{}^{a_3 a_4} + f^{x a_4 y} f_y{}^{a_5 a_3} \right) = 0, \tag{56}$$

but of course not all are independent.

Generally, at n-points each colour factor is an order $n-2$ polynomial in the structure constants f^{abc}. There are $(2n-5)!!$ trivalent diagrams, each of which has $n-3$ internal lines. Each internal line, regarded arbitrarily as say the s-channel, can contribute to one, and only one, colour Jacobi identity with two other diagrams containing the corresponding t- and u-channels. The total number of independent colour Jacobi relations is given by $\sum_{k=1}^{\lfloor (n-2)/2 \rfloor} C_{2k}^{n-2} C_k^{2k} (n-2)!/2^{2k}$. This agrees with the number of independent partial colour-ordered amplitudes. For an n-point tree-level amplitude, the Kleiss–Kuijf relations [167,168] imply that there are at most $(n-2)!$ independent basis partial amplitudes.[14] Using the multi-peripheral colour decomposition of [168] we learn that the number of independent (under the Jacobi identities) colour factors is given by the number of independent (under the Kleiss–Kuijf relations) partial amplitudes, that is $(n-2)!$. Hence, at five points we must have $15 - 6 = 9$ independent Jacobi identities, as claimed.

A set of nine independent Jacobi relations at five points, according to the labelling given in [50], can be chosen as

$$\begin{array}{lll} c_7 = c_6 - c_1 & c_8 = c_2 - c_1 & c_9 = c_3 - c_2; \\ c_{10} = c_4 - c_3 & c_{11} = c_5 - c_4 & c_{12} = c_5 - c_6; \\ c_{13} = c_{10} - c_7 & c_{14} = c_{11} - c_8 & c_{15} = c_{12} - c_9. \end{array} \tag{57}$$

If BCJ duality is valid there should exist a writing of the amplitude such that the nine kinematic identities

$$\begin{array}{lll} n_7 = n_6 - n_1 & n_8 = n_2 - n_1 & n_9 = n_3 - n_2; \\ n_{10} = n_4 - n_3 & n_{11} = n_5 - n_4 & n_{12} = n_5 - n_6; \\ n_{13} = n_{10} - n_7 & n_{14} = n_{11} - n_8 & n_{15} = n_{12} - n_9. \end{array} \tag{58}$$

hold. For any one of the kinematic relations given in (58) it is straightforward to check that it may be satisfied using a slight generalisation of the arguments used in basic

[14] We shall see momentarily that BCJ duality implies that this $(n-2)!$-dimensional basis is over complete, but let us not put the cart before the horse.

four-point example. The challenge is to realise all the relations consistently at once. One route is to first establish some consequences of BCJ duality if it were to hold. Assuming BCJ duality at five points, it was shown in [50] that there must exist a new set relations amongst the $6 = (5-2)!$ partial amplitudes. These were the first example of the *BCJ relations* [50], of which the *fundamental* relations take a particularly simple form,

$$\sum_{i=2}^{n-1} p_1 \cdot \left(\sum_{j=2}^{i} p_j \right) A_n[2, \ldots, i, 1, i+1, \ldots, n] = 0. \tag{59}$$

These relations were established up to eight points explicitly in [50] and conjectured to hold generally on the basis of further significant evidence at higher points. Returning to the 5-point example specifically, and pretending that we do not yet know of the BCJ relations, assuming BCJ duality we can take six numerators $\{n_i\}_{i=1}^{6}$ as independent, the remaining $n_j \notin \{n_i\}_{i=1}^{6}$ being generated by (58). Using the fact that each partial amplitude depends on only five colour-ordered diagrams, we can define, using the kinematic Jacobi relations (58), two of the six n_i, let us say n_5, n_6, in terms of only two partial amplitudes and the remaining four n_i, $i = 1, 2, 3, 4$. For any of the other four independent partial amplitudes, by (58) we can replace any dependence on the n_j not belonging to our choice of independent $\{n_i\}_{i=1}^{6}$ and then further remove any occurrence of n_5, n_6 by their definition in terms of our two special partial amplitudes and the remaining n_i, $i = 1, 2, 3, 4$. Then something unexpected happens - all dependence on $\{n_i\}_{i=1}^{4}$ drops out due to only the kinematic relations amongst the propagators! All six partial amplitudes are simple functions of our two selected partial amplitudes and the propagators alone. These identities generate the KLT and BCJ relations at five points [50].

Now, given the BCJ relations it is possible to explicitly construct a representation of the total amplitude such that BCJ duality holds [52,53]. That this representation yields the correct amplitude is checked via the KLT relations. The loop of reasoning is then cut by demonstrating independently that the BCJ relations do in fact hold. This was done at any number of points in [169,170] by considering the $\alpha' \to 0$ limit of string theory monodromy relations. They may also be deduced from pure spinor cohomology [171]. There are a number of powerful stringy perspectives on the BCJ relations, see for example [172–175]. A purely field theoretic derivation was given in [176] using only Britto-Cachazo-Feng-Witten (BCFW) recursion [177]. They have also been established [178] in $\mathcal{N} = 4$ super Yang–Mills, which contains the pure Yang–Mills case, using the connected formalism of Roiban, Spradlin, Volovich and Witten.

Let us unpack further what we have seen, following closely [179] and the very clear account given in [143]. Consider the n-point tree amplitude written as a sum over $(2m-5)!!$ trivalent graphs. Thinking of the $(2m-5)!!$ colour c_i and kinematic factors n_i as vectors, $\mathbf{c}$, $\mathbf{n}$, we can trivially rewrite the amplitude as

$$\mathcal{A}^{n,0}_{\text{YM}} = \mathbf{c}^t \cdot \mathbf{D} \cdot \mathbf{n} \tag{60}$$

where $[\mathbf{D}]_{ij} = \delta_{ij}/d_j$. Of course, only $(n-2)!$ c_i are independent and we can choose $(n-2)!$ master colour factors and put them into a $(n-2)!$-vector $\mathbf{c}_\mathrm{m}$. The rest are generated by the Jacobi identities

$$\mathbf{c} = \mathbf{J} \cdot \mathbf{c}_\mathrm{m} \tag{61}$$

where $\mathbf{J}$ is a $(2n-5)!! \times (n-2)!$ matrix encoding these relations. For example, at four points, in the conventions of (36), we can choose c_t, c_u as our master colour factors and then

$$\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{62}$$

At five points, for the choice of $\mathbf{c}_\mathrm{m}^t = (c_1, \ldots, c_6)$, in the conventions of (57) we have

$$\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ -1 & 0 & -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 1 & -1 \end{pmatrix}. \tag{63}$$

Of course, the choice of $\mathbf{c}_\mathrm{m}$ is not unique. In this language, BCJ duality amounts to the existence of a rewriting of the amplitude such that

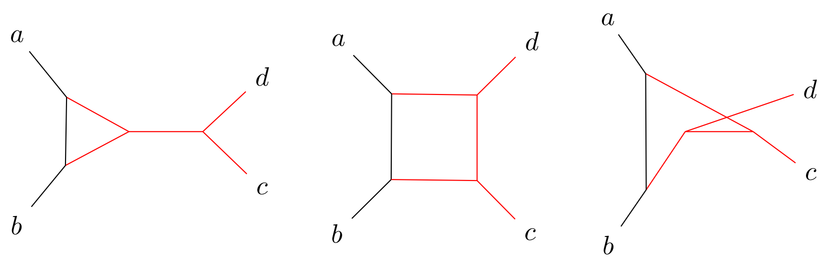
$$\mathbf{n} = \mathbf{J} \cdot \mathbf{n}_\mathrm{m}. \tag{64}$$

But we also have the $(n-2)!$ independent (under Kleiss–Kuijf relations, but prior to applying the BCJ relations) partial amplitudes $\mathbf{A}$, which in this language may be written

$$\mathbf{A} = \mathbf{P} \cdot \mathbf{n} = \mathbf{PJ} \cdot \mathbf{n}_\mathrm{m}. \tag{65}$$

where $\mathbf{P}$ is an $(n-2)! \times (2n-5)!!$ matrix with elements determined by the permutations defining each partial in $\mathbf{A}$ relative to the colour order of the graphs. So the question of identifying a BCJ duality respecting set of numerators reduces to the invertibility of $\mathbf{PJ}$. But wait, what if $\mathbf{PJ}$ is singular? Well that is the point, the BCJ relations imply that only $(n-3)!$ of the partial amplitudes are independent and $\mathbf{PJ}$ *is* singular. But, we *can* solve for $(n-3)!$ elements of $\mathbf{n}_\mathrm{m}$ in terms of $(n-3)!$ partial amplitudes and the remaining $(n-2)! - (n-3)! = (n-3)!(n-3)$ elements of $\mathbf{n}_{\mathrm{m}}$. More generally, any matrix $\mathbf{M}$, with linear system $\mathbf{y} = \mathbf{M} \cdot \mathbf{x}$, admits a generalised inverse $\widetilde{\mathbf{M}}$ satisfying $\mathbf{M}\widetilde{\mathbf{M}}\mathbf{M} = \mathbf{M}$, which implies $\mathbf{y} = \mathbf{M}\widetilde{\mathbf{M}} \cdot \mathbf{y}$. The generalised inverse is not unique, however, so the solution for $\mathbf{n}_{\mathrm{m}}$ given by $\mathbf{n}_{\mathrm{m}} = \widetilde{\mathbf{PJ}} \cdot \mathbf{A}$ is not unique.[15] On substituting this solution back into (65), the dependence of the remaining $(n-3)!(n-3)$ partial amplitudes on $\mathbf{n}_{\mathrm{m}}$ must drop out and we are left with the BCJ relations only; the $(n-3)!(n-3)$ leftover kinematic numerators are entirely unconstrained and may be set to zero at the expense of rendering the surviving kinematic numerators non-local as the propagators corresponding to the vanishing numerators have been shuffled into them.

Our discussion so far has been restricted to tree-level, but to take "gravity $=$ gauge $\times$ gauge" beyond KLT we need to go to loops. The statement of the duality is not affected, up to some minor subtleties that we shall comment on momentarily. For example, consider the simplest example at one loop



(66)

which yields the colour Jacobi identity

$$f^{a}{}_{xa'} f^{xb}{}_{b'} (c_s^{a'b'cd} - c_t^{a'b'cd} - c_s^{a'b'cd}) = 0, \tag{67}$$

where c_s, c_t, c_u are just the four-point tree colour factors given in (43). Under BCJ duality, we should then have

$$n_s - n_t - n_u = 0, \tag{68}$$

but where the kinematic factors are functions of the loop momentum $n = n(\ell)$. The kinematic Jacobi-type identities are functional identities. The four-point one-loop example corresponding to (66) in $\mathcal{N} = 4$ Yang–Mills theory is especially simple, due to the particularly simple structure of one-loop amplitudes [36,180]. See for example [49,144]. For pure Yang–Mills at one and two loops see [181]. For detailed examples at three loops see for example [51,182]. These simple cases make it clear that BCJ can work at loop level. However, the proof of BCJ duality at tree-level given in [52,53] relied on the KLT relations and therefore does not extend to loop level. At the time of writing there is no proof that BCJ duality will hold to all loops, despite an impressive number of highly non-trivial concrete examples [51,57–63,65,67,82,90–93,181,183–185].

15 One could demand that it is a Moore–Penrose pseudo-inverse, which always exists and is unique, picking out one particular solution.

Understanding BCJ duality to all orders is no doubt a central problem, particularly in the context of applications to scattering amplitudes in gravity, the subject of the next section. We shall explore some of the possible routes to BCJ duality later, but let us now summarise some key properties and generalisations of BCJ duality.

2.1.3 BCJ duality: universal properties and generalisations

We have so far only discussed BCJ duality for pure Yang–Mills theory. We did not mention spacetime dimension, so the first obvious generalisation to Yang–Mills theory in arbitrary dimensions is already implicitly contained in our previous discussion. Of course, there are many other possible generalisations that are desirable for a better understanding the principles as well as for applications. For example, to test the UV properties of $\mathcal{N}=8$ supergravity, addressing question (2), we would like to be able to put $\mathcal{N}=4$ Yang–Mills amplitudes into BCJ duality respecting form. More generally, we would like to know what kind of gravity theories can be generated by the BCJ double-copy, addressing question (3), which requires knowing what kind of gauge theories admit a BCJ dual representation. Here we give a lightning tour of the key properties and generalisations:

- *Supersymmetry*: Supersymmetry and BCJ duality are curiously compatible [144]. On the one hand, it is not hard to convince oneself that if BCJ duality holds for pure Yang–Mills, then it will hold for any pure super Yang–Mills multiplet simply through the supersymmetry transformations. Conversely, BCJ duality for Yang–Mills coupled to adjoint fermions implies supersymmetry [73,186]. If we add a single minimally coupled adjoint-valued (minimal) fermion, $\mathcal{L}_{\text{fermion}} \sim \bar{\lambda}\not{D}\lambda$, then BCJ duality requires [73] that for a four-point interaction with only fermions on the external legs we have the following identity,

$$\bar{u}_{[1}\gamma_\mu v_2 \bar{u}_{3]}\gamma^\mu v_4 = 0 \tag{69}$$

  which in turn implies the Fierz identity required for supersymmetry. If BCJ duality is to hold, then supersymmetry follows. Once we have introduced the double-copy we will see that it had to be this way [73].
  As is well-known in the context of supersymmetry, this identity only holds in $D = 3, 4, 6, 10$, so $\mathcal{N} = 1$ Yang–Mills theories only exist in these dimensions [35,36]. Similarly, BCJ duality with a single adjoint fermion only holds in these dimensions. However, BCJ duality survives toroidal dimensional reduction and consistent truncations, so that colour-kinematic duality in $D = 10$, $\mathcal{N} = 1$ implies BCJ duality for all pure super Yang–Mills theories in $D \leq 10$.
  The restriction of $\mathcal{N} = 1$ Yang–Mills theories to $D = 3, 4, 6, 10$ can be related to the existence of the four normed division algebras $\mathbb{A} = \mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ and the fact that they are *alternative* algebras [187–190]. This is reflected by the Lie algebra relation,

$$\mathfrak{so}(1, 1 + \dim \mathbb{A}) \cong \mathfrak{sl}(2, \mathbb{A}) \tag{70}$$

  connecting spacetime and algebra symmetries [191–193]. These structures are in turn related to the notion of triality [194,195] and the triality algebras [196].

Remarkably (or inevitably, depending on your tastes), these structures feed directly into supergravity theories through "gravity = gauge × gauge" with some surprising results [136,138–140,190,197], as we shall describe in section 3.2.2.

- *Bi-adjoint scalars*: The addition of adjoint-valued fermions implied supersymmetry, which typically introduces adjoint-valued scalars, unless it is minimal. In this case, BCJ duality for amplitudes involving scalars is again taken care of by supersymmetry. But can we include adjoint-valued scalars without fermions or supersymmetry? Well, it depends on the couplings. If the scalar field is minimally coupled to the gauge field and has a quartic potential with the Yang–Mills coupling g, $\mathcal{L}_{\phi^4} \sim -g^2\text{tr}([\phi, \phi][\phi, \phi])$, then certainly, since it is merely the dimensional reduction on S^1 of pure Yang–Mills theory in $D+1$. What about a cubic scalar potential with its own coupling constant? The answer is yes, but with the caveat that ϕ must then carry the adjoint representation of a second *global* symmetry group [75]. A gauge invariant cubic term for a set of scalar adjoint scalars $\phi^{\tilde{a}}$, where $\tilde{a}$ is for now an arbitrary global index, may be written,

$$\mathcal{L}_{\phi^3} = \frac{1}{3!}\lambda \tilde{f}_{\tilde{a}\tilde{b}\tilde{c}}\text{tr}([\phi^{\tilde{a}}, \phi^{\tilde{b}}]\phi^{\tilde{c}}) = \frac{1}{3!}\lambda \tilde{f}_{\tilde{a}\tilde{b}\tilde{c}} f_{abc}\phi^{a\tilde{a}}, \phi^{b\tilde{b}}\phi^{c\tilde{c}}. \tag{71}$$

Here, aside necessarily being totally antisymmetric, $\tilde{f}_{\tilde{a}\tilde{b}\tilde{c}}$ is an unconstrained constant tensor. However, for four external scalars at tree-level this term, essentially by construction, contributes to the s-, t- and u-channel kinematic factors one piece of a would-be Jacobi identity each

$$n_s = \lambda^2 \tilde{f}^{\tilde{a}\tilde{b}\tilde{x}} \tilde{f}_{\tilde{x}}{}^{\tilde{c}\tilde{d}} + \cdots, \quad n_t = \lambda^2 \tilde{f}^{\tilde{a}\tilde{x}\tilde{d}} \tilde{f}_{\tilde{x}}{}^{\tilde{b}\tilde{c}} + \cdots, \quad c_u = \lambda^2 \tilde{f}^{\tilde{a}\tilde{x}c} \tilde{f}_{\tilde{x}}{}^{\tilde{d}\tilde{b}} + \cdots. \tag{72}$$

Since these are the unique $\mathcal{O}(\lambda^2)$ contributions, the kinematic Jacobi relation $n_s = n_t + n_u$ (the four-point colour Jacobi relation is clearly unaffected, the only difference being that it is the *scalar* four-point contact term that must the absorbed into the trivalent s, t, u diagrams) requires

$$\tilde{f}^{\tilde{a}[\tilde{b}\tilde{x}} \tilde{f}_{\tilde{x}}{}^{\tilde{c}\tilde{d}]} = 0. \tag{73}$$

Our *a priori* unconstrained tensor $\tilde{f}_{\tilde{a}\tilde{b}\tilde{c}}$ obeys the Jacobi identity and $\phi^{a\tilde{a}}$ is a $G \times \tilde{G}$ bi-adjoint field. For the moment $\tilde{G}$ is just some flavour group, but our suggestive notation is not a coincidence, as we shall see in section 2.2.2. This seems like a rather esoteric addition to the set of fundamental BCJ duality respecting ingredients, especially given that cubic potentials are unbounded from below, but it turns out that bi-adjoint scalar fields are important to the relationship between gauge and gravity amplitudes, as first identified in [198]. They are by now a ubiquitous element of "gravity = gauge × gauge", as evidenced by [75,76,81,86, 109,110,112,113,136,137,179,199–205], and so deserve some special comment.

- *Matter couplings*: We have thus far only considered fields carrying the adjoint representation of the gauge group. What about other representations? Generically, matter fields carrying $\mathfrak{g}$-representations, ρ, ρ' such that $\text{Ad}_{\mathfrak{g}} \in \rho \otimes \rho'$, can couple to the gauge field through the structure constants in the appropriate representation $[T^a]_i{}^j$, where down/up indices belong to the ρ and ρ' representations, respectively. Of course, the properties of the fields involved (commuting vs. Grassmann, spacetime representations etc) will place restrictions on the allowed ρ, ρ'. For quarks, this is the familiar case of $[T^a]_{i\bar{j}} \equiv [T^a]_i{}^j$, where i and $\bar{\imath}$ are fundamental and anti-fundamental representations of SU(3), respectively. So far, so ordinary. But what happens to BCJ duality? In particular, the T's do not satisfy Jacobi relations. Rather, they satisfy commutation relations

$$[T^a]_i{}^j[T^b]_j{}^k - [T^b]_i{}^j[T^a]_j{}^k = f^{ab}{}_d[T^d]_i{}^k. \tag{74}$$

Of course, letting $[f^a]_{bc} = f^a{}_{bc}$ the Jacobi identities are just the commutation relations in the adjoint representation,

$$[f^a]_c{}^d[f^b]_d{}^e - [f^b]_c{}^d[f^a]_d{}^e = f^{ab}{}_d[f^d]_c{}^e, \tag{75}$$

which immediately suggests the appropriate generalisation of BCJ duality to non-adjoint matter [74]: BCJ duality for matter fields is mediated by the commutation relations. For any triple of diagrams involving matter fields with colour factors satisfying a commutation relation, the corresponding kinematic factors must satisfy the same relations. This has been applied to quantum chromodynamics [206] and has applications to black holes physics through the double-copy [130]. Note, quarks (or other matter fields) do not imply supersymmetry, unlike adjoint fermions. Again, in the context of the double-copy this is perfectly natural as we shall come to discuss. Note that for particular choices of gauge group and matter representations there may be additional colour identities, beyond the basic commutation relations, specific to these choices [74]. However, as emphasised in [74], these need not be imposed; what is essential to BCJ duality are generic identities that do not hinge on any special properties of the representations or gauge groups used.

- *Symmetry breaking*: We previously mentioned consistent truncations of pure super Yang–Mills theories as a method for producing BCJ duality respecting theories with less supersymmetry. This is the simplest example of a broader class of symmetry breaking principles that preserve BCJ duality at tree-level. These are particularly useful in the construction of large classes of supergravity theories using the double-copy [73,75,76,86,87,87,88,207]. We shall discuss some of these in subsequent sections.
  One can both spontaneously and explicitly break symmetries while preserving the BCJ relations [76]. Consider a subgroup $H \subset G$ corresponding to the positive eigenspace subspace of a Cartan involution $\theta : \mathfrak{g} \to \mathfrak{g}$. The adjoint representation decomposes as $\mathrm{Ad}_G = \mathrm{Ad}_H \oplus \rho_H$, where ρ_H is a (not necessarily irreducible) representation of H. Under this subalgebra any adjoint-valued multiplet φ_{Ad_G} will decompose accordingly,

$$\varphi_{\mathrm{Ad}_G} \to \varphi_{\mathrm{Ad}_H} \oplus \varphi_{\rho_H}, \tag{76}$$

  where φ_{Ad_H} and φ_{ρ_H} belong to the positive and negative eigenspaces of θ, respectively. If φ_{Ad_G} transforms under some further global symmetry group $\mathcal{G}$, which

may include R-symmetry, we may also decompose to a subgroup $\mathcal{H} \subset \mathcal{G}$. Combining both the explicit gauge and global group breakings, we can effect various truncations and introduce matter representations in manner that automatically preserves BCJ duality at tree-level. Particularly useful examples are given by field theory orbifolds [73]. Take a some order-k element $\tau = \rho_{\mathcal{G}}(\sigma) \otimes \mathrm{Ad}_G(g)$, and project the fields φ onto the τ-invariant subsector. The perhaps simplest example is given by $\tau := (-1)^F \cdot \theta$, where F is the fermion number operator. In this case, for adjoint-valued bosons b and fermions f, the τ-invariant subsector is given by $b_{\mathrm{Ad}_H}, f_{\rho_H}$.
Hence, we generate matter representations f_{ρ_H} starting from purely adjoint fields in such a way that BCJ duality is inherited. This clearly generalises to supermultiplets. For example, following the same procedure we can project an adjoint $D = 4, \mathcal{N} = 4$ supervector multiplet onto an adjoint $D = 4, \mathcal{N} = 2$ supervector multiplet plus a fundamantal $D = 4, \mathcal{N} = 2$ hyper multiplet.

- *Other algebraic structures*: The presentation thus far would understandably leave one with the impression that trivalent diagrams are essential. However, this is really an artefact of the fact that Lie algebras have rank-3 structure constants. It is the algebraic structure of the gauge symmetry that dictates the nature of the duality. Given a generalised gauge theory with something other than a Lie algebra underpinning its colour structure, the BCJ duality will reflect its fundamental identities, which need not be the Jacobi relation. Although this principle is reasonable, working examples are rare. An important case is given in [71], where $\mathcal{N} = 16$ supergravity [208,209] was derived from the double-copy of the $D = 3, \mathcal{N} = 8$ Bagger–Lambert–Gustavsson theory [210–212] through a colour-kinematic duality based on the fundamental Lie three algebra identity. Since $\mathcal{N} = 16$ supergravity is the unique maximally supersymmetric theory in $D = 3$, this construction must be equivalent to that of the standard BCJ double-copy of $D = 3, \mathcal{N} = 8$ Yang–Mills and indeed it is [70]. Further related examples were explored in [213], including the Aharony–Bergman–Jafferis–Maldacena theory [214] although the BCJ relations are absent beyond six points in that case. Where could one search for generalised gauge theories admitting a novel colour-kinematic duality structure? Since the Bagger–Lambert–Gustavsson theory is also a higher gauge theory [215] this could be one broad avenue to explore.
- There are various geometric or world-sheet perspectives on BCJ duality and "gravity = gauge × gauge". For example, the BCJ relations and duality can be cast in string theoretic terms [53,169,170,174,216–218]. There are also "world-sheet theories for field theory", in particular the scattering equation formalism [199,201,219,220] and the ambi-twistor string approach [221–225], which provides a route to a better all-loop understanding through a "nodal" world-sheet genus expansion.

2.2 The Bern–Carrasco–Johansson double-copy construction

The notion of BCJ duality is an intriguing property of gauge theories in its own right, with non-trivial and otherwise hidden implications, such as the reduction of the number of independent partial amplitudes down to $(n-3)!$. It is not yet clear why it should hold and there are many open questions left to explore, in particular whether or not it can be made manifest and, relatedly, if it holds to all loops. However, the (remarkably sheltered!) reader might be forgiven for asking what this all has to do with "gravity = gauge × gauge". Here is the answer: given BCJ duality holds for pure Yang–Mills then every $\mathcal{N}=0$ supergravity amplitude follows directly from gluon amplitudes through the *BCJ double-copy* [51,64]. Note, this statement only depends on the validity of the BCJ colour-kinematic conjecture, otherwise it is completely generic; it applies to all non-Abelian gauge groups in all spacetime dimensions. In this section we shall describe this construction. We begin with a review of perturbative $\mathcal{N}=0$ supergravity, before exploring its double-copy construction. Finally, we shall layout the growing zoology of double-copy constructible theories and discuss their implications for perturbative quantum gravity.

2.2.1 $\mathcal{N}=0$ supergravity

The common, or NS-NS, sector of the $\alpha' \to 0$ limit of closed string theories is given by $\mathcal{N}=0$ supergravity,

$$S_{\mathcal{N}=0} = \frac{1}{2\kappa^2}\int \star R - \frac{1}{(D-2)} d\varphi \wedge \star d\varphi - \frac{1}{2} e^{-\frac{4}{D-2}\varphi} H \wedge \star H, \tag{77}$$

where $2\kappa^2 = 16\pi G_{\mathrm{N}}^{(D)}$. Aside from the metric g we have the dilaton φ and the KR 2-form $H = dB$. The solutions of the associated equations of motions give backgrounds (with vanishing cosmological constant) around which strings can be quantised to lowest order in α' and string coupling. They also ensure conformal invariance of the string is non-anomalous in critical dimensions. From the relationship between the double-copy and the KLT relations, we should not be surprised by the appearance of this action in "gravity = gauge × gauge".

In $D=4$, for topologically trivial manifolds, we can dualise the B into a pseudo-scalar, the axion χ. To do so, forget B and consider H as the field with respect to which we vary. Of course, H is closed and so we must add a Lagrange multiplier χ to enforce this condition

$$\mathcal{L} = \star R - \frac{1}{2} d\varphi \wedge \star d\varphi - \frac{1}{2} e^{-2\varphi} H \wedge \star H - d\chi \wedge H. \tag{78}$$

Varying with respect to χ we find $dH = 0$ as required. Now, varying with respect to H we find

$$H = -\frac{1}{2} e^{2\varphi} \star d\chi. \tag{79}$$

Since this is algebraic we substitute back into $\mathcal{L}$ to obtain

$$\mathcal{L}_{\text{dual}} = \star R - \frac{1}{2} d\varphi \wedge \star d\varphi - \frac{1}{2} e^{2\varphi} d\chi \wedge \star d\chi, \tag{80}$$

which is often referred to as axion–dilaton gravity. This is equivalent (semi-classically, at least [226]) to $D = 4, \mathcal{N} = 0$ supergravity. We emphasise the dual axion–dilaton picture as it highlights a rather general feature of simple double-copy constructible theories. To describe this we need one further step. Consider the $\mathfrak{sl}(2, \mathbb{R})$ generators

$$E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{81}$$

$$[H, E_\pm] = \pm 2E_\pm, \quad [E_+, E_-] = H. \tag{82}$$

Then $\mathcal{V} = e^{\frac{1}{2}\varphi} e^{E_+\chi}$ is an SL$(2, \mathbb{R})/$SO(2) coset representative in the "positive root gauge" and

$$\frac{1}{4}\mathrm{tr}\left(d\mathcal{M}^{-1} \wedge \star d\mathcal{M}\right) = -\frac{1}{2}d\varphi \wedge \star d\varphi - \frac{1}{2}e^{2\varphi}d\chi \wedge \star d\chi, \tag{83}$$

where $\mathcal{M} = \mathcal{V}^T\mathcal{V}$. This makes the invariance of $\mathcal{L}_{\text{dual}}$ under global SL$(2, \mathbb{R})$ transformations $\mathcal{V} \mapsto \mathcal{V}M$, $\det(M) = 1$, and local (in the sense that they are functions of φ, χ) SO(2) transformations, $\mathcal{V} \mapsto M(\varphi, \chi)\mathcal{V}$, $M(\varphi, \chi)^T M(\varphi, \chi) = 1$, manifest. A *symmetric* homogeneous space $\mathcal{G}/\mathcal{H}$ satisfies

$$[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h}, \tag{84}$$

where $\mathfrak{g} = \mathfrak{h} + \mathfrak{p}$. From the commutation relations (82) we note that SL$(2, \mathbb{R})/$SO(2) is symmetric. For a complete characterisation of symmetric space see, for example, [227,228].

The fact that the scalars parametrise a symmetric space is an almost generic property of double-copy constructible gravity theories. We say "almost generic", as there are numerous exceptions [77,85,229], but it is true for all the basic examples, of which there are many. As we shall describe, it is possible to understand when such scalar manifolds appear from "gravity = gauge × gauge" and why they should be consistent with the double-copy construction [72,85,138]. Despite this, there is no complete proof that the double-copy yields a symmetric spaces when it should, although the statement has passed all tests at the level of symmetries and amplitudes to date.

Let us now turn to scattering amplitudes. Ignoring φ, B for now, we can expand the Einstein–Hilbert action perturbatively around a Minkowski background $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$

$$S_{\text{EH}} \sim \int d^D x \sum_{n=0} \kappa^n h^{n+1} \Box h \tag{85}$$

and construct amplitudes as pioneered by Bryce DeWitt [230–232]. The Feynman diagrams for gravitons include n-point vertices for all n. However, just as for the four-point vertex in Yang–Mills, these can all be absorbed into the kinematic numerators of the purely trivalent diagrams. For example, consider as in [143] a pure cubic diagram i, contributing N_i/d_i, where N_i is the kinematic factor (no colour factors here as we are dealing with gravitons), to the amplitude integrand, and another diagram $i_{(4)}$, contributing $N_{i_{(4)}}/d_{i_{(4)}}$, which is identical except that one cubic four-point sub-diagram with propagator s has been contracted to a four-point vertex. Then $d_{i_{(4)}} = d_i/s$ and so

$$\frac{N_i}{d_i} + \frac{N_{i_{(4)}}}{d_{i_{(4)}}} = \frac{N_i + sN_{i_{(4)}}}{d_i} = \frac{N'_i}{d_i}. \tag{86}$$

Having written the graviton amplitude in terms of pure cubic diagrams, it takes a form resembling closely the gluon amplitude (33),

$$A^{n,L}_{g,B,\varphi} = i^L \left(\frac{\kappa}{2}\right)^{n-2+2L} \sum_i \int \prod_{l=1}^{L} \frac{d^D p_l}{(2\pi)^D S_i} \frac{N_i}{d_i}. \tag{87}$$

2.2.2 The double-copy

Although not obvious, for factorisable external states, which form a basis, the gravitational kinematic numerators can always be written as a product $N_i = n_i \tilde{n}_i$. This brings us to the statement of the BCJ double-copy prescription for pure Yang–Mills [51,64]:

BCJ double-copy prescription: Given any two pure Yang–Mills n-point L-loop amplitudes

$$\begin{aligned} \mathcal{A}^{n,L}_{\mathrm{YM}} &= i^L g^{n-2+2L} \sum_i \int \prod_{l=1}^{L} \frac{d^D p_l}{(2\pi)^D S_i} \frac{c_i n_i}{d_i}, \\ \widetilde{\mathcal{A}}^{n,L}_{\mathrm{YM}} &= i^L g^{n-2+2L} \sum_i \int \prod_{l=1}^{L} \frac{d^D p_l}{(2\pi)^D S_i} \frac{c_i \tilde{n}_i}{d_i}, \end{aligned} \tag{88}$$

at least one of which respects BCJ colour-kinematic duality[16], let us assume it is $\mathcal{A}^{n,L}_{\mathrm{YM}}$, we may "double-copy" by replacing the colour factors by BCJ respecting kinematic factors while sending $g \to \kappa/2$, to generate a new amplitude,

$$A^{n,L}_{g,B,\varphi} = i^L \left(\frac{\kappa}{2}\right)^{n-2+2L} \sum_i \int \prod_{l=1}^{L} \frac{d^D p_l}{(2\pi)^D S_i} \frac{n_i \tilde{n}_i}{d_i}, \tag{89}$$

which is guaranteed to be a *bona-fide* amplitude of $\mathcal{N} = 0$ supergravity.

Some immediate comments are in order:

1. The external states of $A^{n,L}_{g,B,\varphi}$ are determined by the tensor product of the external states of $\mathcal{A}^{n,L}_{\mathrm{YM}}$ and $\widetilde{\mathcal{A}}^{n,L}_{\mathrm{YM}}$, which need not be the same. The external states are

[16] The c_i should not be explicitly evaluated under the integral in case they accidentally vanish before being replaced by the loop-momenta dependent kinematic factors.

labelled by their on-shell spacetime little group representations SO($D-2$) and, more generally, any other global representation they carry. For example, in $D=4$ Yang–Mills we have the various possible products of gluons with ± 1 helicity states:

$$\begin{array}{c|c|c} \otimes & +1 & -1 \\ \hline +1 & +2,\ \text{graviton} & 0,\ \tau \\ +1 & 0,\ \bar{\tau} & -2,\ \text{graviton} \end{array} \tag{88}$$

where $\tau = \varphi + e^{i\chi}$. Of course we get the $2 \times 2 = 4$ degrees of freedom of $\mathcal{N}=0$ supergravity on a Minkowski background.

2. The amplitudes $\mathcal{A}^{n,L}_{\text{YM}}$ and $\widetilde{A}^{n,L}_{\text{YM}}$ need not derive from the same theory, as long as both theories admit a BCJ duality respecting form. This allows one to construct the product of different theories. The spectrum of states of the gravity theory is given by tensor product of the left and right gauge theory factors. For example, if one factor is $\mathcal{N}=4$ Yang–Mills and the other is pure $\mathcal{N}=2$ Yang–Mills then the double-copy is $\mathcal{N}=6$ supergravity.
   Varying over all BCJ compatible factors we generate a panoply of double-copy constructible gravity theories. Clearly, if one wishes to restrict to a single graviton, then each factor must have at most one massless adjoint gauge field. Note, however, the left and right factor need not have any gauge fields at all. For example, the amplitudes of $\mathcal{N}=2$ hyper multiplet amplitudes generate those of $\mathcal{N}=4$ Maxwell theory: "gauge = matter × matter". However, for the hyper multiplets to have a local symmetry they must come coupled to an $\mathcal{N}=2$ Yang–Mills multiplet, which will generate the $\mathcal{N}=4$ gravitational sector when included in the double-copy. So the $\mathcal{N}=4$ Maxwell amplitudes generated by the hypers must be regarded as a subsector of a double-copy theory including the gravitational degrees of freedom.
3. Invariance of the gauge theory amplitudes under the linearised gauge transformations together with BCJ duality implies the invariance of the double-copy amplitudes under linearised diffeomorphisms and hence that they belong to some gravitational theory [86]. The emergence of linearised diffeomorphism can also be seen directly at the level of field theory [137], as we shall discuss in Sect. 3. The same field theoretic mechanism generates a linearised local supersymmetry for every adjoint fermion belonging to the factors [137] and so their product should be locally supersymmetric, i.e. a (possibly generalised) supergravity theory, where the gravitini follow from the product of the adjoint gluons and fermions. But if the product is locally supersymmetric the factors must be globally supersymmetric, which is precisely consistent with the observation that BCJ duality and adjoint fermions together implies supersymmetry. Conversely, if one or both of the factors have global supersymmetry, then the corresponding invariance of the gauge amplitudes, together with BCJ duality, implies that the amplitudes of the double-copy theory are invariant under linearised local supersymmetry transformations [144].
4. Of course, we can always go back to one of the gauge amplitudes by turning kinematics back into colour. We can proceed further and replace the remaining kinematics by a second copy of the colour, leaving us with an amplitude of the bi-adjoint ϕ^3 theory introduced in Sect. 2.1.3. This reflects the idea expressed in [198] that the relationship between gauge theory and gravity is heuristically of the

form "ϕ^3 × gravity = gauge × gauge". For tree-level Yang–Mills at four points, imposing BCJ duality, this is quite literally the case,

$$A^{4,0}_{\varphi^3} A^{4,0}_{g,B,\varphi} = \widetilde{A}^{4,0}_{\mathrm{YM}} \widetilde{A}^{4,0}_{\mathrm{YM}}. \tag{89}$$

It is also manifest at the level of integrands in the scattering equation formalism for any number of points [199]. There is another interpretation of the form "gravity = gauge $\times \tilde{\phi}^3 \times$ gauge", where $\tilde{\phi}$ is in some sense the inverse of ϕ [109,136,137,199]. At tree-level this can be made concrete using double-partial Yang–Mills amplitudes

$$A^{n,0}_{g,B,\varphi} = \mathbf{A}^{n,0\,t}_{\mathrm{YM}} \cdot \mathbf{S}_{\mathrm{KLT}} \cdot \mathbf{A}^{n,0}_{\mathrm{YM}}. \tag{90}$$

Here, $\mathbf{A}^{n,0}_{\mathrm{YM}}$ is a specific choice of $(n-3)!$ independent partial amplitudes and $\mathbf{S}_{\mathrm{KLT}}$ is the corresponding momentum kernel [233], which in this case is exactly the inverse of the matrix of scalar propagators, that is the double-partial amplitudes of the ϕ^3 theory [199].

This picture allows one to construct a vast array of gravitational theories in the sense that every double-copy of a pair a gauge theory amplitudes gives a gravitational amplitude of some theory and that every amplitude of that theory can be written as a double-copy.

2.2.3 A growing zoology

The basic principle is that if one can cast two gauge theories into BCJ duality respecting form, then their amplitudes yield a double-copy theory with spectra given by the tensor product of the spectra of the two gauge theories: double-copy states = left gauge states ⊗ right gauge states. There is a growing list of BCJ compatible gauge theories and, thus, double-copy constructible theories. Gauge invariance in the left and right factors implies linearised diffeomorphisms in the double-copy theory, so it will (typically) be a theory of gravity. Here we summarise the known double-copy constructible theories together with their gauge factors and the key ideas facilitating the implementation of BCJ duality and identifying the double-copy theory. Rather than follow the chronology, we will start with the simplest examples and work up in complexity.[17] We are not consistent in our labelling of the classes of double-copies as it is easiest to characterise them in terms of the factors in some cases and the double-copy theory in others. In all cases we are strictly referring only to the tree-level theories, although in many examples they have passed numerous loop-level tests.

1. ϕ^3 **theory:** The very simplest thing one can do is to replace not the colour factors, but the kinematic factors of pure Yang–Mills theory:

$$\frac{n_i c_i}{d_i} \to \frac{\tilde{c}_i c_i}{d_i}. \tag{91}$$

[17] We are measuring complexity in terms of how simple it is to fully characterise the space of theories of interest: $\mathcal{N}=8$ supergravity is unique and determined by supersymmetry alone, as simple as can be, whereas the space of non-supersymmetric theories containing Einstein–Hilbert gravity is wild, even with reasonable consistency conditions imposed.

It is not hard to convince oneself that this yields the amplitudes of the bi-adjoint ϕ^3 theory, with interaction (71), as identified in [199].

2. **Pure $\mathcal{N}$-extended Yang–Mills** $\times$ **pure $\tilde{\mathcal{N}}$-extended Yang–Mills**: As we have mentioned adjoint fermions and BCJ duality implies supersymmetry. Conversely, given BCJ duality for gluons, supersymmetry extends it to any pure super Yang–Mills multiplet. Consequently, the double-copy of super Yang–Mills theories follows directly. Of course, the left and right vectors $A, \tilde{A}$ alone yield $\mathcal{N} = 0$ supergravity, but the vectors with the $\mathcal{N}, \tilde{\mathcal{N}}$ left/right adjoint fermions yields $\mathcal{N}+\tilde{\mathcal{N}}$ gravitini. Hence, the double-copy must be a supergravity theory. For $\mathcal{N} + \tilde{\mathcal{N}}$ half-maximal or greater, there is unique candidate supergravity theory so the identification of the double-copy theory is trivial; it is just read-off from the tensor product of the left/right gauge theories states given in Table 2. For less supersymmetry, the couplings are not uniquely determined by the spectra so there is *a priori* some ambiguity. This can be resolved by examining the symmetries alone. This was done in [138,140] for all pure super Yang–Mills theories in $D = 3, \ldots 10$, as reviewed in Sect. 3.2.2. The key observation is that the scalar manifolds of the resulting supergravity theories are symmetric homogeneous spaces $\mathcal{G}/\mathcal{H}$, as can be deduced by consistently truncating the maximally supersymmetric examples. This in turn implies that the Lagrangian is uniquely determined by the non-compact global symmetry group $\mathcal{G}$. The BCJ double-copy has been explicitly established and tested at the loop level for many of the cases [56,57,91,95,104,183,184]. An important approach to the double-copy construction of such theories (and others beside) is through orbifoldings of $\mathcal{N} = 8$ supergravity that factorise into orbifolds of the left and right $\mathcal{N} = 4$ Yang–Mills theories [68,73]. We summarise results in Table 2. For our conventions and details of the various supermultiplets see Table 1. For the remaining discussion we will mostly focus on $D = 4$ and simply comment on other dimensions.

3. **$\mathcal{N} = 4$ and $\mathcal{N} = 3$ supergravity coupled to vector multiplets**: Restricting to $D = 4$ the next simplest class of double-copy constructible supergravity theories is $\mathcal{N} = 4$ supergravity coupled to an arbitrary number of vector multiplets [68,72,73]. From the above list we see that $\mathcal{N} = 2 \times \tilde{\mathcal{N}} = 2$ yields $\mathbf{V}_2 \otimes \tilde{\mathbf{V}}_2 = \mathbf{G}_4 \oplus 2\mathbf{V}_4$ so that we have at least two vector multiplets. If one restricts to adjoint-valued multiplets then this is the only consistent $\mathcal{N} = 2 \times \tilde{\mathcal{N}} = 2$ case; adding extra vector or hyper multiplets into the factors would generate additional graviton and/or gravitini multiplets. Of course, there is the other possibility of $\mathcal{N} = 4 \times \tilde{\mathcal{N}} = 0$. This yields pure $\mathcal{N} = 0$ supergravity [183]. However, one can couple n adjoint-valued scalars to the right factors (with couplings determined from the dimensional reduction of pure Yang–Mills in $D = 5$) to give $\mathcal{N} = 4$ supergravity coupled to n vector multiplets, $\mathbf{V}_4 \otimes [\tilde{A} \oplus n\phi] = \mathbf{G}_4 \oplus n\mathbf{V}_4$, with global symmetry $\mathrm{SL}(2, \mathbb{R}) \times \mathrm{SO}(6, n)$ [67,68,72,90–92,138,183]. This can be trivially extended to half-maximal supergravities in $D = 3, 5, 6, 7, 8, 9, 10$. This exhausts all half-maximal or greater supergravity theories: the double-copy spans all such theories.
The story for $\mathcal{N} = 3$ is slightly different. There is no perturbative $\mathcal{N} = 3$ super Yang–Mills theory, since $\mathcal{N} = 3$ supersymmetry implies $\mathcal{N} = 4$ for Yang–

Table 1 On-shell helicity states of all $D = 4$ supermultiplets

Q	R	Multiplet$_{\mathcal{N}}$	f	$\text{Un}(1)_{st} \times R$ representations
32	SU(8)	$\mathbf{G}_8$	256	$\mathbf{1_{-4} + 8_{-3} + 28_{-2} + 56_{-1} + 70_0 + \overline{56}_1 + \overline{28}_2 + \overline{8}_3 + 1_4}$
28	Un(7)	$\mathbf{G}_7$	256	$\mathbf{1^0_{-4} + 7^1_{-3} + 1^{-7}_{-3} + 21^2_{-2} + 7^{-6}_{-2} + 35^3_{-1} + 21^{-5}_{-1} + \overline{35}^4_0 + c.c.}$
24	Un(6)	$\mathbf{G}_6$	128	$\mathbf{1^0_{-4} + 6^1_{-3} + 15^2_{-2} + 1^{-6}_{-2} + 20^3_{-1} + 6^{-5}_{-1} + \overline{15}^4_0 + c.c.}$
20	Un(5)	$\mathbf{G}_5$	64	$\mathbf{1^0_{-4} + 5^1_{-3} + 10^2_{-2} + \overline{10}^3_{-1} + 1^{-5}_{-1} + \overline{5}^4_0 + c.c.}$
16	Un(4)	$\mathbf{G}_4$	32	$\mathbf{1^0_{-4} + 4^1_{-3} + 6^2_{-2} + \overline{4}^3_{-1} + 1^4_0 + c.c.}$
16	SU(4)	$\mathbf{V}_4$	16	$\mathbf{1_{-2} + 4_{-1} + 6_0 + \overline{4}_1 + 1_2}$
12	Un(3)	$\mathbf{G}_3$	16	$\mathbf{1^0_{-4} + 3^1_{-3} + \overline{3}^2_{-2} + 1^3_{-1} + c.c.}$
12	Un(3)	$\mathbf{V}_3$	16	$\mathbf{1^0_{-2} + 3^1_{-1} + 1^{-3}_{-1} + \overline{3}^2_0 + c.c.}$
8	Un(2)	$\mathbf{G}_2$	8	$\mathbf{1^0_{-4} + 2^1_{-3} + 1^2_{-2} + c.c.}$
8	Un(2)	$\mathbf{V}_2$	8	$\mathbf{1^0_{-2} + 2^1_{-1} + 1^2_0 + c.c.}$
8	Un(2)	$\mathbf{H}_2$	8	$\mathbf{1^r_{-1} + 2^{r+1}_0 + 1^{r+2}_1 + c.c.}$
8	Un(2)	$\frac{1}{2}\mathbf{H}_2$	4	$\mathbf{1^{-1}_{-1} + 2^0_0 + 1^1_1}$
4	Un(1)	$\mathbf{G}_1$	4	$(-4, 0) + (-3, 1) + c.c.$
4	Un(1)	$\mathbf{V}_1$	4	$(-2, 0) + (-1, 1) + c.c.$
4	Un(1)	$\mathbf{C}_1$	4	$(-1, r) + (0, r + 1) + c.c.$
0	-//-	A	2	$(-2) + c.c.$
0	-//-	λ	2	$(-1) + c.c.$
0	-//-	ϕ	2	(0)

Here Q counts the number of supercharges, R denotes the global R-symmetry group, **Multiplet**$_{\mathcal{N}}$ the type of $\mathcal{N}$-extended supermultiplet and f is number of degrees of freedom. The $\mathcal{N}$-extended gravity, vector, hyper and chiral multiplets are denoted by $\mathbf{G}_{\mathcal{N}}$, $\mathbf{V}_{\mathcal{N}}$, $\mathbf{H}_{\mathcal{N}}$ and $\mathbf{C}_{\mathcal{N}}$, respectively. Note, $\frac{1}{2}\mathbf{H}_2$ is used to denote a half-hyper multiplet. Although $\mathbf{V}_3$ and $\mathbf{V}_4$ are identical as isolated gauge multiplets, when coupled to supergravity they must be distinguished. Similarly, $\mathbf{G}_7$ and $\mathbf{G}_8$ have identical content and as interacting theories are identical despite having *a priori* distinct symmetries. Finally, we use A, λ and ϕ to denote the smallest $\mathcal{N} = 0$ vector, spinor and scalar multiplets, respectively. Sub/superscripts in the final column refer to the Un(1) charges carried by the representations. The subscripts refer to the spacetime little group $\text{Un}(1)_{\text{st}}$ helicities, which we uniformly multiply by a factor of two to make the notation more compact. The superscripts refer to the internal Un(1) charges. When the symmetry has no semi-simple part we use tuplets $(a, b, c, \ldots)$ to label the Un(1) charges, with the first slot reserved for $\text{Un}(1)_{\text{st}}$

Mills.[18] Hence, there is only one way to obtain $\mathcal{N} = 3$ through the double-copy, $\mathbf{V}_2 \otimes \tilde{\mathbf{V}}_1 = \mathbf{G}_3 \oplus \mathbf{V}_3$. The double-copy necessarily comes coupled to at least one vector multiplet. Adding adjoint-valued multiplets (vector, hyper or chiral) to either factor would result in extra graviton or gravitini multiplets so is forbidden (without increasing the degree of supersymmetry) (Table 2). However, we are able to include hyper and chiral multiplets in non-adjoint representations, using the matter (by which we mean any fields not valued in the adjoint) colour-kinematic duality of [74,206]. Let us include a single half-hyper multiplet in the left $\mathcal{N} = 2$ factor in a pseudo-real representation ρ (required for half-hypers), which is com-

[18] But not supergravity, the analog in this case is $\mathcal{N} = 7$ supersymmetry implies $\mathcal{N} = 8$.

Table 2 Summary of supergravity theories obtained from the double-copy of two pure super Yang–Mills theories

D	$\mathcal{N}\otimes\tilde{\mathcal{N}}$	Double-copy	$\mathcal{G}$	Comments
3	$\mathbf{V}_8\otimes\tilde{\mathbf{V}}_8$	$\mathbf{G}_{16}$	$E_{8(8)}$	BLG double-copy [70]
	$\mathbf{V}_8\otimes\tilde{\mathbf{V}}4$	$\mathbf{G}_{12}$	$E_{7(-5)}$	
	$\mathbf{V}_8\otimes\tilde{\mathbf{V}}2$	$\mathbf{G}_{10}$	$E_{6(-14)}$	
	$\mathbf{V}_8\otimes\tilde{\mathbf{V}}1$	$\mathbf{G}_9$	$F_{4(-20)}$	
	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}4$	$\mathbf{G}_8\oplus 4\mathbf{V}_8$	SO(8, 4)	
	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}2$	$\mathbf{G}_6\oplus 2\mathbf{V}_6$	SU(4, 2)	
	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}1$	$\mathbf{G}_5\oplus 2\mathbf{V}_5$	USp(4, 2)	
	$\mathbf{V}_2\otimes\tilde{\mathbf{V}}2$	$\mathbf{G}_4\oplus 2\mathbf{V}_4$	SU(2, 1) × SU(2, 1)	
	$\mathbf{V}_2\otimes\tilde{\mathbf{V}}1$	$\mathbf{G}_3\oplus 1\mathbf{V}_3$	SU(2, 1)	
	$\mathbf{V}_1\otimes\tilde{\mathbf{V}}1$	$\mathbf{G}_2\oplus 1\mathbf{V}_2$	SL(2, $\mathbb{R}$)	
4	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}_4$	$\mathbf{G}_8$	$E_{7(7)}$	4-point 5-loop finite [95]
	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}_2$	$\mathbf{G}_6$	SO*(12)	4-point 2-loop amplitudes [184]
	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}_1$	$\mathbf{G}_5$	SU(5, 1)	Enhanced cancellations [57]
	$\mathbf{V}_2\otimes\tilde{\mathbf{V}}_2$	$\mathbf{G}_4\oplus 2\mathbf{V}_4$	SL(2, $\mathbb{R}$) × SO(6, 2)	
	$\mathbf{V}_2\otimes\tilde{\mathbf{V}}_1$	$\mathbf{G}_3\oplus 1\mathbf{V}_3$	Un(3, 1)	
	$\mathbf{V}_1\otimes\tilde{\mathbf{V}}_1$	$\mathbf{G}_2\oplus 1\mathbf{H}_2$	Un(2, 1)	
5	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}_4$	$\mathbf{G}_8$	$E_{6(6)}$	
	$\mathbf{V}_4\otimes\tilde{\mathbf{V}}_2$	$\mathbf{G}_6$	SU*(6)	
	$\mathbf{V}_2\otimes\tilde{\mathbf{V}}_2$	$\mathbf{G}_2$	O(5, 1)O(1, 1)	
6	$\mathbf{V}_{(1,1)}\otimes\tilde{\mathbf{V}}_{(1,1)}$	$\mathbf{G}_{2,2}$	SO(5, 5)	
	$\mathbf{V}_{(1,1)}\otimes\tilde{\mathbf{V}}_{(0,1)}$	$\mathbf{G}_{2,1}$	SU*(4) × USp(2)	
	$\mathbf{V}_{(1,0)}\otimes\tilde{\mathbf{V}}_{(0,1)}$	$\mathbf{G}_{1,1}$	O(4) × O(1, 1)	
7	$\mathbf{V}_1\otimes\tilde{\mathbf{V}}_1$	$\mathbf{G}_2$	SL(5, $\mathbb{R}$)	
8	$\mathbf{V}_1\otimes\tilde{\mathbf{V}}_1$	$\mathbf{G}_2$	SL(2, $\mathbb{R}$) × SL(3, $\mathbb{R}$)	
9	$\mathbf{V}_1\otimes\tilde{\mathbf{V}}_1$	$\mathbf{G}_2$	SL(2, $\mathbb{R}$) × O(1, 1)	
10	$\mathbf{V}_{(1,0)}\otimes\tilde{\mathbf{V}}_{(0,1)}$	$\mathbf{G}_{(1,1)}$	O(1, 1)	Type IIA
	$\mathbf{V}_{(1,0)}\otimes\tilde{\mathbf{V}}_{(1,0)}$	$\mathbf{G}_{(2,0)}$	SL(2, $\mathbb{R}$)	Type IIB

patible with BCJ duality [77], and n chiral multiplets on the $\tilde{\mathcal{N}}=1$ right theory, also in a pseudo-real representation $\tilde{\rho}$:

$$[\mathbf{V}_2\oplus\tfrac{1}{2}\mathbf{H}_2^{\rho}]\otimes[\mathbf{V}_1\oplus\mathbf{C}_2^{\tilde{\rho}}]=\mathbf{G}_3\oplus(n+1)\mathbf{V}_3. \tag{92}$$

For convenience the tensor product of the on-shell matter multiplet are given in Table 3. Note, the "matter" representations $\rho,\tilde{\rho}$ do not double-copy with the adjoint-valued fields since the amplitudes necessarily have distinct colour structures and the Jacobi identities are replaced with commutation relations [74]. As in the case of $\mathcal{N}=4$ supergravity the vector multiplet coupling is unique [234] and

Table 3 The content $D = 4$ resulting from the product of the on-shell helicity states of left and right matter multiplets, as summarised in Table 1

$\mathcal{N}$	$\otimes$	$\tilde{\mathcal{N}}$	Result
$\mathbf{H}_2$	$\otimes$	$\tilde{\mathbf{H}}_2$	$4\mathbf{V}_4$
$\mathbf{H}_2$	$\otimes$	$\tilde{\mathbf{C}}_1$	$2\mathbf{V}_3$
$\mathbf{H}_2$	$\otimes$	$\tilde{\lambda}$	$2\mathbf{V}_2$
$\frac{1}{2}\mathbf{H}_2$	$\otimes$	$\tilde{\mathbf{H}}_2$	$\mathbf{V}_4$
$\frac{1}{2}\mathbf{H}_3$	$\otimes$	$\tilde{\mathbf{C}}_1$	$\mathbf{V}_3$
$\frac{1}{2}\mathbf{H}_2$	$\otimes$	$\tilde{\lambda}$	$\mathbf{V}_2$
$\mathbf{C}_1$	$\otimes$	$\tilde{\mathbf{C}}_1$	$\mathbf{V}_2 \oplus \mathbf{H}_2$
$\mathbf{C}_1$	$\otimes$	$\tilde{\lambda}$	$\mathbf{V}_1 \oplus \mathbf{C}_1$

the scalars belong to the symmetric spaces

$$\frac{\mathrm{SU}(3, 1+n)}{\mathrm{SU}(3) \times \mathrm{Un}(1+n)}. \tag{93}$$

Through dimensional reduction/oxidation this exhausts the analysis for all Poincaré supergravity theories with more than eight supercharges ($\mathcal{N} > 2$ in $D = 4$). Every such theory is double-copy constructible with the single exception of pure $D = 4, \mathcal{N} = 3$ supergravity and its dimensional reductions/oxidations.

4. **$\mathcal{N} = 2$ supergravity with homogeneous scalar manifolds:** The complete classification for more than eight supercharges relied on the fact that the scalar manifolds of supergravity are, in this case, necessarily symmetric homogenous spaces. For eight (or fewer) supercharges, there is far more freedom. The scalar manifolds are required to be *special geometries* [235,236], which includes real, Kähler and quarternionic manifolds [237–239], but homogeneity is not essential. Consequently the space of theories is far richer in this case.
Focussing on $D = 4$, the scalars belonging to vector multiplets must parametrise a projective special Kähler manifold [235,238,240], while those belong to hyper multiplets parametrise a quarternionic-Kähler manifold [241–243]. A manageable, in the sense that there is an explicit and complete characterisation, subclass of $\mathcal{N} = 2$ supergravity theories is given by those with homogenous scalar manifolds. A unified double-copy construction of almost all $\mathcal{N} = 2$ supergravity theories coupled to vector multiplets with homogenous scalar manifolds was given in [77] through a left $\mathcal{N} = 2$ Yang–Mills theory coupled to a single half-hyper multiplet in a pseudo-real representation and right $\tilde{\mathcal{N}} = 0$ Yang–Mills theory coupled to adjoint scalars and pseudo-real fermions. If non-symmetric the scalar manifolds are indexed by three integers $(q, P, \dot{P})$,

$$\mathrm{SO}(1,1) \times \frac{\mathrm{SO}(q+2,2)}{\mathrm{SO}(q+2) \times \mathrm{Un}(1)} \times \frac{\mathcal{S}_q(P, \dot{P})}{S_q(P, \dot{P})} \ltimes \left[(\mathbf{spin}, \mathbf{def}, \mathbf{1})^1 \ltimes (\mathbf{1}, \mathbf{1}, \mathbf{1})^2\right], \tag{94}$$

where **spin** indicates the spinor representation of $SO(q+2, 2)$ and **def** the defining representation of $\mathcal{S}_q(P, \dot{P})$. Here, $(q, P, \dot{P})$ are integers, which fix the number of

vector multiplets, the factor $S_q(P, \dot{P})$ and representations carried by the fields. See, for example, [85] for full details.
If the scalar manifold is symmetric there are three classes: (i) the generic Jordan sequence [244] indexed by a single integer, $(q, P, \dot{P}) = (q, 0, 0)$, (ii) the four magic supergravities [237,244,245] for which $(q, P, \dot{P}) = (n, 1, 0)$, where $n = \dim \mathbb{A} = 1, 2, 4, 8$, and (iii) the minimally coupled sequence [246] indexed by a single integer, $(q, P, \dot{P}) = (-2, P, 0)$. The scalar manifolds are respectively

$$\frac{\mathrm{SU}(1,1)}{\mathrm{Un}(1)_g} \times \frac{\mathrm{SO}(q+2,2)}{\mathrm{SO}(q+2)\times U(1)}; \quad \frac{\mathrm{Conf}(\mathfrak{J}_3^{\mathbb{A}})}{[\mathrm{Str}_0(\mathfrak{J}_3^{\mathbb{A}_\mathbb{C}})]_c}; \quad \frac{\mathrm{SU}(1,P+1)}{\mathrm{Un}(1)\times \mathrm{SU}(P+1)}, \tag{95}$$

where $\mathbb{A}_\mathbb{C} \cong \mathbb{C} \otimes \mathbb{A}$, $\mathfrak{J}_3^{\mathbb{A}}$ is the cubic Jordan algebra of 3×3 Hermitian matrices over $\mathbb{A} = \mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ and $\mathfrak{J}_3^{\mathbb{A}_\mathbb{C}} \cong \mathbb{C} \otimes \mathfrak{J}_3^{\mathbb{A}}$ its complexification, $\mathrm{Conf}(\mathfrak{J})$ is the conformal group of the cubic Jordan algebra $\mathfrak{J}$, $\mathrm{Str}_0(\mathfrak{J})$ is the reduced structure group and $[G]_c$ denotes the compact real form of the complexified group G. The minimally coupled sequence was given as a truncation of the generic Jordan sequence, but can also be constructed directly [85]. This list includes almost all $\mathcal{N} = 2$ supergravity theories coupled to vector multiplet with scalars parametrising a homogenous manifold. The only exceptions are pure $\mathcal{N} = 2$ supergravity and the T^3 model,[19] which cannot be double-copy constructed [85].[20] For recent work on the double-copy construction of this class of theories at one-loop see [229]. One can in principle include an arbitrary number of hype rmultiplets with homogenous scalars manifolds [85], completing the classification of this subclass of double-copy constructible theories, although in the non-symmetric case it is not clear how it is to be realised. All cases may be summarised by

$$\begin{aligned} &\mathbf{G}_2 \oplus (1+q+2+r)\mathbf{V}_2 \oplus (q'+4+t/2)\mathbf{H}_2 \\ &\quad = \left[\mathbf{V}_2 \oplus \mathbf{H}_2^{\rho}\right] \otimes \left[\tilde{V} \oplus (q+2)\tilde{\phi} \oplus (r)\tilde{\lambda}^{\tilde{\rho}} \oplus 2(q'+4)\Phi^{\tilde{\rho}} \oplus (t)\varphi^{\tilde{\rho}}\right], \end{aligned} \tag{96}$$

where the specific supergravity theory obtained is detemined by the various parameters and choices of couplings and symmetries of the right gauge theory, as described in detail in [77,85].

5. **Einstein–Maxwell–Yang–Mills supergravity:** Thus far all vector multiplets appearing in the double-copy theory have been Abelian. It possible to also introduce Yang–Mills multiplets through a simple mechanism [75]. The left theory is take to be a pure super Yang–Mills theory, while the right theory is given by pure Yang–Mills coupled to a $\tilde{G} \times G'$ bi-adjoint scalar $\phi^{aa'}$, where $\tilde{G}$ is the gauge group

[19] $\mathcal{N} = 2$ supergravity coupled to a single vector multiplet with non-compact global symmetry group $\mathrm{SL}(2, \mathbb{R})$, under which the two Maxwell field strengths and their duals transform as the **4** [246,247]. Can be regarded as the "symmetrisation" of the STU model [248], where the there complex scalars S, T, U are identified.

[20] But see final point of this list.

as usual, but G' is a global symmetry,

$$\tilde{\mathcal{L}}_{\mathrm{YM}+\phi^3} = \mathrm{tr}\left(\frac{1}{2}F\wedge\star F + \frac{1}{2}D\phi^{a'}\wedge\star D\phi^{a'} - \frac{g^2}{4}\star[\phi^{a'},\phi^{b'}][\phi^{a'},\phi^{b'}] - \frac{g\lambda}{3!}\star f_{a'b'c'}\phi^{a'}[\phi^{b'},\phi^{c'}]\right). \tag{97}$$

The key observation is that the global symmetry G' of the right theory is promoted to a gauge symmetry of the corresponding vector multiplets, $\mathbf{V}_{\mathcal{N}}\otimes\phi^{a'}$, of the double-copy theory with coupling determined by the cubic scalar term of (97), $g'\sim\kappa\lambda$ [75]. Combined with the previous techniques this allows for the double-copy construction of a variety of $\mathcal{N}\leq 4$ Einstein–Maxwell–Yang–Mills supergravity theories [75]. There have been several subsequent tree-level [86,249–251], one-loop [252,253] and all-loop for a single external graviton [86] developments of Einstein–Maxwell–Yang–Mills amplitudes. Finally, the double-copy Einstein–Maxwell–Yang–Mills supergravity theories can be Higgsed by taking the left factor on the Coulomb branch and introducing matching masses for the scalars on the right through an explicit symmetry breaking [76].

6. **Gauged Poincaré supergravity:** Gauged supergravity theories with Minkowski background can also be constructed [87]. Here a subgroup of the R-symmetry of the corresponding ungauged supergravity is gauged, leading to massive gravitini. The left theory is a Higgsed Yang–Mills theory coupled to a set of scalars, which introduces the required massive bosons. The right theory super Yang–Mills theory is also Higgsed and has explicitly broken (through orbifolding) supersymmetry, which introduces the required has massive fermions to generate massive gravitini. Starting with an $\mathcal{N}=2$ super Yang–Mills theory, the simplest examples generate $\mathrm{Un}(1)$ gaugings of the generic Jordan supergavities discussed above [87]. However, it is possible to extended to more supersymmetry and non-Abelian gaugings [87, 88].
7. **$\mathcal{N}\leq 1$ (super)gravity:** For theories with less supersymmetry it is much harder to make general statements. Of course, there is the central example of $\mathcal{N}=0$ supergravity, but beyond it is difficult to characterise what classes of gravity theories may be double-copy constructed. This is principally due to the lack of symmetry, which makes it more difficult to identify what theory is generated by the double-copy. Nonetheless, there are numerous examples. This includes orbifoldings of "parent" double-copy supergravity theories that preserve BCJ duality, but break all, or almost all, the supersymmetries [73]. Control over the resulting theories follows from the control over the parent theory and the relevant orbifold. This technique has, for example, been used to construct all "twin" supergravity theories in [81].[21] This procedure generates new theories from old such as, for example, the $\mathcal{N}=1$ twins in $D=4$. Again, control over the nature of these theories is inherited from their parents. There are various examples, both at tree-level and for loops [73,81,86], but a coherent picture remains to be developed. A particularly

[21] Twin supergravities are pairs of supergravity theories with identical bosonic sectors but distinct supersymmetric completions [81,245,254–258].

important example is the double-copy of Yang–Mills coupled to quarks [74,206]. As well as being of interest in its own right, the techniques developed in this context have opened the door to a vastly expanded array of doubly-copy theories, as the preceding discussion of examples that make use of non-adjoint representations makes clear.

8. **Conformal (super)gravity**: In cases previously considered here, the graviton sector has been Einstein–Hilbert. Remarkably, however, this is not necessary. A counter example is given by conformal (super)gravity [83,259]. The key idea is to use in one of the gauge theory factors a higher derivative $(DF)^2$ theory. In conjunction with various deformations one can then double-copy construct a number of conformal (super)gravity theories, including the Berkovits-Witten theory [260] and (mass-deformed) minimal conformal supergravity. See [261] for a review of conformal supergravity.
9. **Exceptions to the exceptions:** We should be clear about our definition of a "double-copy constructible" theory. The double-copy theory is defined by the *totality* generated by the two gauge theory factors: a particular theory is double-copy constructible if (1) all its amplitudes can be generated by the double-copy of the amplitudes two BCJ duality[22] respecting theories and conversely (2) all amplitudes of the two theories generate an amplitude belonging to the corresponding gravitational theory.[23] For example, a conspicuous absentee is good old Einstein–Hilbert gravity; it is *not* double-copy constructible in the above sense, as it always comes with the axion–dilaton sector. Similarly, the T^3 model is not double-copy constructible [85], which rather stands out as the only case of an $\mathcal{N} \geq 2$ supergravity theory with a (non-trivial) symmetric scalar manifold not admitting a double-copy construction.
   However, as always there are exceptions to the exceptions. All amplitudes of pure Einstein–Hilbert gravity *can* be systematically double-copy constructed by consistently restricting the external states to the graviton sector, while cancelling the would-be axion–dilaton sector appearing in loops with the product of "ghost" chiral fermion amplitudes [74]. The restriction on the external states violates our strong definition, but all amplitudes of Einstein–Hilbert gravity may nonetheless be double-copy constructed using these "ghost" cancellations. With this understanding of "double-copy constructible" it *may* be possible to fill in all the gaps, as well providing alternative constructions of double-copy theories. For example, pure $\mathcal{N} = 4$ supergravity may also be constructed through pure $\mathcal{N} = 2$ Yang–Mills $\times\tilde{\mathcal{N}} = 2$ Yang–Mills using ghost cancellations to remove the unwanted vector multiplet [82].

The above list is by no means exhaustive, although it clearly demonstrates the long-arm of the double-copy construction. In particular, we have not discussed the double-copy construction of: open and closed string amplitudes using Z-theory [79,80,89,262,263]; Born-Dirac-Infield theories, including couplings to super Yang–

[22] This may be trivially true, for example a cubic theory of scalars transforming in the adjoint of a global symmetry.

[23] As it stands this can of course only be established in general at tree-level, with supporting evidence from case-by-case examples of loop-level amplitudes. Our present analysis is explicitly tree-level only.

Mills theories and non-linear sigma models [86,201,264]; the special-Galileon theory [204]; and double-copy correlator relations [265–267]. It is also possible to apply the "gravity $=$ gauge $\times$ gauge" perspective to construct, discover, or deduce properties of, theories for which there is no, and perhaps can be no, Lagrangian description in the conventional sense [66,138,142,258,268,269]. For example, previously unknown $D=4, \mathcal{N}=2$ superconformal S-fold theories of the type introduced in [270–273], which being intrinsically non-perturbative are not amenable to the double-copy proper, where discovered using this approach in [258]. Finally, an alternative and elegant realisation, at tree-level, of many of these "gravity $=$ gauge $\times$ gauge" examples, and the relations between, them is given by the Cachazo-He-Yuan scattering equation formalism [199,201,219]. The amplitudes in this framework can be regarded as a world-sheet integrals, but localised on the solutions of the scattering equations. They sit in-between a string and particle picture. This formalism may also be derived from ambi-twistor string theory [221,274], which then opens a route to loops [223,224,275] and curved backgrounds [225]. It has also been employed to construct candidate tree-level amplitudes for the $D=6, \mathcal{N}=(4,0)$ theory, conjectured to arise in M-theory [276], through the double-copy of $(2,0)$ theory amplitudes [269]. This is all the more remarkable in light of the fact that we, at present, have no other insight regarding the interacting $(4,0)$ theory.

3 Field theory relations

These developments raise the question: to what extent, or in what sense, can one regard gravity as the square of Yang–Mills. Is there a deeper connection underlying the amplitude relations. Having exposed the hidden dualities of amplitudes through an intrinsically on-shell window, is it possible to now step back and understand their origins from a geometric or off-shell point of view. This is not only a conceptual question; having an off-shell understanding may shine light on the outstanding amplitude questions, such as BCJ duality beyond tree-level. There are a number of approaches one might take: can BCJ duality be manifested at the level of the Lagrangian or field equations [200,204,277–282]; can we rewrite the gravity in a form that, in some sense, factorises [64,204,281,283–285]; is there a field theory "product" of gauge theories [81,85,115,116,119,136–142]. We can also turn this on-shell versus off-shell question around: can the BCJ double-copy paradigm be repurposed to efficiently construct solutions in theories of gravity from gauge amplitudes and/or solutions [109–135]. In a sense this runs contrary to the "on-shell paradigm" that took us here. Going back off-shell may nonetheless be instructive.

3.1 Manifesting BCJ duality and the double-copy off-shell

3.1.1 BCJ Lagrangians and kinematic algebras

The remarkable relationship between colour and kinematics hidden in the amplitudes suggests that there is some underlying kinematic algebra mirroring the properties of conventional Lie algebras [200,277,278,280,282]. In general, the nature of this conjectured hidden algebra is not known, however for the self-dual sector it can be identified precisely as an area-preserving diffeomorphism Lie algebra in a particular two-dimensional subspace [277]. To see this, recall the self-duality constraint, $F_{\mu\nu} = \frac{i}{2}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$, reduces the equations of motion in light-cone gauge $A_u = 0$ to

$$\Box\varphi + ig[\partial_w\varphi, \partial_u\varphi] = 0, \tag{98}$$

where $u = t - z$, $v = t - z$, $w = x + iy$ and $A_w = 0$, $A_v = -\frac{1}{4}\partial_w\varphi$, $A_{\overline{w}} = -\frac{1}{4}\partial_u\varphi$. For suitable boundary conditions (98) can be solved perturbatively in momentum space $\varphi^a(p) = \sum_n \varphi^a_n(p)$. Schematically, we have

$$\begin{aligned}\varphi^a_n(p) = \frac{1}{2}g^n \sum_i \int \prod_{l=1}^{n+1} \frac{dp_l}{(2\pi)^4} \left(\frac{n_i(F)^p{}_{p_1 p_2 \ldots p_{n+1}} c_i(f)^a{}_{a_1 a_2 \ldots a_{n+1}}}{p^2 d_i} \right) \\ \times \varphi_0^{a_1}(p_1)\varphi_0^{a_2}(p_2)\cdots\varphi_0^{a_{n+1}}(p_{n+1}).\end{aligned} \tag{99}$$

Each order n correction can be represented in terms of $(n + 1)$-point trivalent tree diagrams, labelled here by i, with n sources of momenta p_l and one external field $\varphi^a_n(p)$. The d_i term appearing denominator is the product of the momenta squared of the internal lines. The important components from our perspective are the kinematic and colour numerators $n_i(F)$ and $c_i(f)$. As usual $c_i(f)$ is a polynomial in the gauge group structure constants f^{abc} generated by attaching one to each vertex. Remarkably, $n_i(F)$ is constructed in precisely the same way with f^{abc} replaced by a kinematic "structure constant" $F^{p_a p_b p_c}$ such that $c_i \leftrightarrow n_i$ under $f \leftrightarrow F$. Specifically, $F_{p_1 p_2}{}^q = (2\pi)^4\delta^4(p_1 + p_2)(p_{1w}p_{2u} - p_{1u}p_{2w})$, where indices are raised/lowered by $\delta^{pq} = \delta_{pq} = (2\pi)^4\delta^4(p + q)$ with contractions given by integration $X_{p\ldots}Y^{p\cdots} := \int \frac{dp}{(2\pi)^4} X(p,\cdots)Y(p,\cdots)$. Using these conventions $F^{p_a p_b p_c}$ is totally anti-symmetric and obeys the Jacobi identity [277], which combined with $c_i \leftrightarrow n_i$ under $f \leftrightarrow F$ makes the BCJ colour-kinematic duality manifest at the level of perturbative classical solutions in the self-dual sector. The kinematic structure constants F are those of the algebra of infinitesimal area-preserving diffeomorphisms. Moreover, this algebra has been shown to determine the kinematic numerators of tree-level maximally helicity violating amplitudes in the complete Yang–Mills theory including the anti-self-dual sector [277]. Understanding these structures beyond the self-dual sector remains an important open question.

Another approach is to modify the Yang–Mills action so that it manifests the duality between colour and kinematics directly in its Feynman diagrams [64]. One can in principle constructively determine the BCJ duality respecting Lagrangian order-by-order [279],

$$\mathcal{L}_{\text{BCJ}} = \mathcal{L}_{\text{YM}} + \mathcal{L}_{(5)} + \mathcal{L}_{(6)} + \cdots . \tag{100}$$

Of course, $\mathcal{L}_{(n)}$ are constrained to leave the amplitudes invariant, but nevertheless rearrange the kinematic numerators. This was done explicitly in [64] to six points. For example, choosing Feynman gauge one possibility at five points is given by

$$\mathcal{L}_{(5)} \propto f_{[a_1 a_2}{}^{b} f_{a_3]b}{}^{c} f_{c a_4 a_5} \partial_{[\mu} A^{a_1}_{\nu]} A^{a_2}_{\rho} A^{a_3 \mu} \frac{1}{\Box} (A^{a_4 \nu} A^{a_5 \rho}). \tag{101}$$

This is identically zero since we have the Jacobi identity sitting up front $f_{[a_1 a_2}{}^{b} f_{a_3]b}{}^{c} = 0$, hence the amplitudes are trivially left invariant by the addition of $\mathcal{L}_{(5)}$. However, separating the three terms, trivially inserting $\Box/\Box$ for each term, and redistributing them, as we would in the amplitude, shifts the numerators of the five-point diagrams such that they are in BCJ dual form [64]. Let us assume we had found BCJ numerators starting from the original Lagrangian. Adding $\mathcal{L}_{(5)}$ would preserve the duality by construction; that we can add identically zero terms to the Lagrangian while maintaining BCJ duality is another way to see that the BCJ numerators are non-unique. Note, BCJ duality can be made completely manifest at the level of the Lagrangian for a non-linear sigma model [204]. The BCJ double-copy of the non-linear sigma model yields the special Galileon and "squaring" the non-linear sigma model action gives a novel form of Galileon action [204].

This brings us to the idea that the $\mathcal{N} = 0$ supergravity Lagrangian can be "factorised". What does one mean by this? In the context of string the left and right movers heuristically give rise to the spacetime indices on $h_{\mu\nu}, B_{\mu\nu}, \varphi$. Thus, the left/right indices of $Z_{\mu\nu} \sim h_{\mu\nu} + B_{\mu\nu} \sim A_\mu \tilde{A}_\nu$ have their origin in the left/right open strings corresponding to the left/right gauge theories. Given that each gauge theory is independent, one might therefore anticipate a formulation of $\mathcal{N} = 0$ supergravity that makes this manifest in that the left and right indices only "talk" amogst themselves. It is in this sense that we mean the action factorises. This idea was sometime ago proposed by Siegel, who demonstrated that there does indeed exist such a formalism, at least for specific gauge choices [286,287]. Later Grant and Bern developed a perturbative Lagrangian that manifests the left/right split order-by-order [283]. To give a simple illustration of the idea Bern and Grant imposed de Donder gauge and made a field redefinition for the metric perturbation $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$ and the dilaton φ

$$h_{\mu\nu} \to h_{\mu\nu} + \eta_{\mu\nu}\sqrt{\frac{2}{D-2}}\varphi, \quad \varphi \to \tfrac{1}{2}h + \sqrt{\frac{D-2}{2}}\varphi, \tag{102}$$

which yields at zeroth order

$$\mathcal{L}_{\text{EH}} = -\tfrac{1}{2} h^{\mu}{}_{\nu} \Box h_{\mu}{}^{\nu} + \varphi \Box \varphi. \tag{103}$$

The terms contracting amongst the "left" and "right" indices (an ambiguous notion since $h_{\mu\nu}$ is symmetric) have been removed. Of course, this condition has to be maintained to all orders. The field redefinition realising this goal, even before making any gauge choice, is remarkably simple,

$$g_{\mu\nu} = e^{\sqrt{\frac{2}{D-2}}\kappa\varphi} e^{\kappa h_{\mu\nu}}, \quad \varphi \to \sqrt{\frac{2}{D-2}} \left(\varphi + \frac{1}{2}h\right), \tag{104}$$

and was checked explicitly through order six in κ, allowing the KLT relations to be derived directly from the action itself up to five points [283].

Here the dilaton was introduced as an auxiliary device to aid the factorisation of the Einstein–Hilbert Lagrangian, but given the nature of the double-copy one should only expect the full factorisation to work for $\mathcal{N} = 0$ supergravity, where the dilaton and KR 2-form are genuine components of the full theory.[24] Indeed, one could take the view that the non-symmetric $Z_{\mu\nu}$ is required to make sense of the notion of having left and right indices at all. This was taken seriously in [284], where a left/right factorised action was constructed using the double-field theory formalism [288–290], which enlarges the set of spacetime coordinates to accommodate a (symmetric) generalised metric to render the dualities of string theory manifest. The generalised metric was introduced in the context of string and membrane dualities in earlier related work [286,287,291–294]. As particularly relevant to "gravity = gauge × gauge", the generalised metric of [291] was obtained in [286] using left and right vierbeins, making the left/right sectors apparent with a manifest $\mathrm{GL}(D, \mathbb{R}) \times \mathrm{GL}(D, \mathbb{R})$ symmetry. More recent approaches to this question [281,285] have also made a twofold Lorentz symmetry (rather than $\mathrm{GL}(D, \mathbb{R})$, since only the metric was considered) manifest to all orders [281]. This formulation also has the potentially appealing feature, from the BCJ double-copy point of view, that the left/right factorised Lagrangian of [285] has only cubic interactions with the aid of only a single auxiliary field $a^{\rho}{}_{\mu\nu}$,

$$\mathcal{L}_{\mathrm{EH}} \sim a^{\rho}{}_{\mu\nu}\partial_{\rho}\mathfrak{g}^{\mu\nu} - \left(a^{\rho}{}_{\sigma\mu}a^{\sigma}{}_{\rho\nu} - \frac{1}{D-1}a^{\rho}{}_{\rho\mu}a^{\sigma}{}_{\sigma\nu}\right)\mathfrak{g}^{\mu\nu} \tag{105}$$

where $\mathfrak{g}^{\mu\nu}$ is the usual tensor density $\sqrt{-g}g^{\mu\nu}$.

3.1.2 Double-copy solutions

One can apply the BCJ double-copy paradigm to the construction of classical solutions in theories of gravity, such as black holes, from gauge theory. This may take the guise of applying a classical double-copy-like map to classical gauge theory solutions or extracting perturbative classical solutions from the double-copy of gauge theory amplitudes [109–135].

Let us consider the former case. In its simplest incarnation, introduced in [109], there is double-copy-like map of gauge theory solutions that yields non-perturbative solutions in Einstein–Hilbert gravity under the assumption that the spacetime metric is of Kerr–Schild type,

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \phi(x)k_{\mu}k_{\nu}, \tag{106}$$

where $\phi(x)$ is, morally speaking, related to the by-now familiar bi-adjoint scalar, although it carries no indices here. The covector field k_{μ} is null with respect to both g and η. The Minkowski background can be generalised to an arbitrary background, in which case k_{μ} is null and geodesic with respect to the background metric. The Kerr–Schild form of the metric effectively linearises the Ricci tensor.

[24] Of course, for tree-level amplitudes they may be consistently truncated to leave only the graviton scattering amplitudes of Einstein–Hilbert gravity.

To give a feeling for the classical double-copy let us turn to the simplest example:

$$\text{Schwarschild black hole} = (\text{static colour charge})^2. \tag{107}$$

More specifically, the solution to the sourced Yang–Mills equation $D\star F = j$, where $j_a^\mu = -gc_a\delta(\mathbf{x})(1, 0, 0, 0)$ is a static point-like colour charge located at the origin with constant c_a, is taken to be

$$A_\mu^a = \phi c^a k_\mu, \quad k_\mu = (1, \hat{\mathbf{x}}), \quad \phi = \frac{1}{4\pi r} \tag{108}$$

which obviously linearises the Yang–Mills equation. Now, in precise analogy to the BCJ double-copy we send the gauge coupling g to the gravitational coupling $\kappa/2$ and the colour factor c^a to a second copy of the kinematics $c^a \mapsto Mk_\nu$, where M is a mass-dimension-one constant. The scalar ϕ goes along for the ride, just as for the propagators in the BCJ double-copy . It is in this sense that it is related to ϕ^3-theory. Hence

$$A_\mu^a = \phi c^a k_\mu \mapsto \frac{\kappa}{2}\frac{M}{4\pi r}k_\mu k_\nu \tag{109}$$

which we recognise as nothing but the Schwarschild solution in Kerr–Schild coordinates

$$g_{\mu\nu}^{\text{Schwar.}}(x) = \eta_{\mu\nu} + \frac{\kappa}{2}\frac{M}{4\pi r}k_\mu k_\nu = \eta_{\mu\nu} + \frac{2GM}{r}k_\mu k_\nu \tag{110}$$

with static point-like mass located at the origin $T_{\mu\nu} = Mv_\mu v_\nu\delta(\mathbf{x})$, where $v^\mu = (1, 0, 0, 0)$, which is the obvious "double-copy" of j^μ. The solution (108) is perhaps a little unfamiliar, but is related by a gauge transformation to the standard Coulomb solution $A_\mu^a = \frac{gc^a}{4\pi r}(1, 0, 0, 0)$. This serves to highlight the subtle role played by gauge and coordinate choices in the context of solutions, as opposed to amplitudes [117,119,132]. Indeed, making another gauge choice for the point charge solution it is possible instead to obtain a gravity solution including a dilatonic contribution [114,117]; it would seem the that the Kerr–Schild classical double-copy is not unique [117]. In fact, the most general Kerr–Schild classical double-copy of the Coulomb solution has been argued perturbatively [117] to be the two-parameter Janis–Newman–Winicour solution [295], which can be tuned to turn off the dilaton, leaving the Schwarschild solution. The exact Kerr–Schild double-copy realisation of this space of solutions was given in [296], which, interestingly, used the T-duality generalised metric [297] and double field theory formalism [288], generalising the class of solutions considered in the Kerr–Schild double-field theory double-copy of [298].

This basic example has since been extended to a number of (generalised) Kerr–Schild spacetimes [109–116,118–125,296,298]. It is also possible to construct spacetimes perturbatively using a direct classical analog of the BCJ duality and the double-copy [117].

3.2 Field theoretic "gravity = gauge × gauge"

Another approach to addressing such questions is to build a dictionary at the level of fields, as opposed to on-shell states or amplitudes, expressing the covariant fields of (super)gravity in terms of the product of (super) Yang–Mills fields. That is, can we interpret

$$A_\mu(x)\text{``}\otimes\text{''}\tilde{A}_\nu(x). \tag{111}$$

directly in the context of field theory, without appealing to on-shell conditions? Invoking the known properties of open and closed strings we can deduce a consistent indentification of the product of gauge potentials or vector supermultiplets [299,300]. But is there an independent definition of "⊗" at the level of field theory which is valid whether or not there is an underlying string interpretation to guide our identifications? This raises two immediate sub-questions: (i) gravity has no colour, so where do the left and right gauge groups go? (ii) amplitudes are multiplicative in momentum space; is this reflected in the product? Said another way, does the product violate the Leibnitz rule?

From the Weinberg–Witten theorem the product cannot be a straightforward tensor product of any kind. Moreover, the lessons of the double-copy strongly suggest a subtle relationship with the ϕ^3-theory. It should at least be compatible with the intricacies of BCJ duality and the double-copy in cases where the product field theory agrees with that generated by the double-copy of the corresponding amplitudes of the factor theories.

Guided by the structure of the amplitude relations and requirements of symmetry a covariant product rule was introduced in [137]. It is independent of the amplitude relations, but, in all cases where we have been able to test it, it is compatible in the sense that product field theory agrees with the amplitude product. It is defined as:

$$f \circ \tilde{f} := \langle\langle f \cdot \Phi \cdot \tilde{f}\rangle\rangle. \tag{112}$$

Here, f, $\tilde{f}$ are arbitrary spacetime fields valued in $\mathfrak{g}$ and $\tilde{\mathfrak{g}}$, respectively. The "spectator" field $\Phi = \Phi^{a\tilde{a}} T_a \otimes \tilde{T}_{\tilde{a}}$ is a $G \times \tilde{G}$ bi-adjoint valued scalar. The $\cdot$ product denotes an associative convolutive inner tensor product with respect to the Poincaré group

$$[f \cdot g](x) = \int d^D y f(y) \otimes g(x-y) \tag{113}$$

and $\langle\langle\ ,\ ,\ \rangle\rangle : \mathfrak{g} \times (\mathfrak{g} \otimes \tilde{\mathfrak{g}}) \times \tilde{\mathfrak{g}} \to \mathbb{R}$ is a trilinear trace form constructed from the negative-definite trace forms of $\mathfrak{g}$, $\tilde{\mathfrak{g}}$, which in the standard basis is simply,

$$\langle\langle X, Y, Z\rangle\rangle = X_a Y^{a\tilde{a}} Z_{\tilde{a}}. \tag{114}$$

The convolution reflects the fact that the amplitude relations are multiplicative in momentum space. For sufficiently well-behaved functions the convolution obeys,

$$\partial_\mu [f \cdot g](x) = [\partial_\mu f \cdot g](x) = [f \cdot \partial_\mu g](x). \tag{115}$$

This turns out to be essential for reproducing the local symmetries of (super)gravity from those of the two (super) Yang–Mills factors. The double trace form accounts for the gauge groups, while the spectator field allows for arbitrary and independent G and $\tilde{G}$. Of course, it is closely related to the bi-adjoint scalar of the BCJ zeroth-copy. Heuristically, it can be considered as its convolutive pseudo-inverse $\Phi \sim \phi^{-1}$.

3.2.1 Local symmetries

Having introduced the covariant product, let us consider the case of two pure Yang–Mills theories. The field-theoretic product of two gauge potentials, A_μ and $\tilde{A}_\nu$, is given by

$$[A_\mu \circ \tilde{A}_\nu](x) = g^2[A^a_\mu \cdot \Phi_{a\tilde{a}} \cdot \tilde{A}^{\tilde{a}}_\nu](x). \tag{116}$$

In addition to the gauge potential, one must also include the accompanying BRST ghost fields [137,141,142]. This reflects the fact that we must include the gauge transformations while the product is defined on fields. Thus it is natural to include the BRST ghosts into the "gravity = gauge × gauge" construction. Indeed, the inclusion of BRST ghosts in the context of Yang–Mills squared or "open × open" strings was advocated sometime ago by Siegel [299,300]. With the ghosts incorporated, the total product of left and right pure Yang–Mills theories is given schematically by

$\circ$	$\tilde{A}_\nu$	$\tilde{c}^\beta$
A_μ	$g_{\mu\nu}$ (graviton) $+$ $B_{\mu\nu}$ (KR 2-form) $+\eta_{\mu\nu}\varphi$ (dilaton)	$\tilde{C}^\beta_\mu$ (right diffeo. + KR ghosts)
c^α	C^α_ν (left diffeo. + KR ghosts)	$\lambda^{(\alpha\beta)}$ (KR ghost-for-ghosts) $+ (g+\varphi)\varepsilon^{\alpha\beta}$ (det g + dilaton)

(117)

Here we have introduced the $SL(2,\mathbb{R})$-doublets of left/right ghosts and anti-ghosts, $c^\alpha = (c, \bar{c})$, following [301]. This dictionary is heuristic, we give the precise relationship below at the linear level, but lays out the basic structure. First, it splits into four sectors:

(i) $A \times \tilde{A} =$ physical + auxiliary
(ii) $A \times \tilde{c} =$ right ghosts
(iii) $c \times \tilde{A} =$ left ghosts
(iv) $c \times \tilde{c} =$ ghosts-for-ghosts + physical/auxiliary

This is quite intuitive, except perhaps for the mixing of physical and auxiliary degrees of freedom in the $A \times \tilde{A}$ and $c \times \tilde{c}$ sectors. This mixing is a consequence of choosing Einstein frame, as opposed to string frame. We will make this precise momentarily. Second, the ghost numbers and mass dimensions of the $\mathcal{N} = 0$ supergravity follow consistently from the product. First, ghost numbers $\mathrm{gh}(f)$ and Grassmann grades $\varepsilon(f)$ are additive under the product

$$\begin{aligned} \mathrm{gh}(f \circ g) &= \mathrm{gh}(f) + \mathrm{gh}(g); \\ \varepsilon(f \circ g) &= \varepsilon(f) + \varepsilon(g) \mod 2; \end{aligned} \tag{118}$$

Similarly, since the mass dimension of the spectator[25] is $(3D+2)/2$ we have

$$[f\circ g]=[f]+[g]-\frac{D-2}{2} \tag{119}$$

which as we shall see is precisely as required. The ghost number, grade and mass dimension, $(\mathrm{gh}(f), \varepsilon(f), [f])$, of the product are summarised here:

$$\begin{array}{c|c|cc} \circ & \begin{array}{c}\tilde{A}^{\tilde{a}}_{\nu}\\ (0,0,\frac{D-2}{2})\end{array} & \begin{array}{c}\tilde{c}^{\tilde{a}}\\ (1,1,\frac{D-2}{2})\end{array} & \begin{array}{c}\tilde{\bar{c}}^{\tilde{a}}\\ (-1,1,\frac{D-2}{2})\end{array} \\ \hline \begin{array}{c}A^{a}_{\nu}\\ (0,0,\frac{D-2}{2})\end{array} & (0,0,\frac{D-2}{2}) & (1,1,\frac{D-2}{2}) & (-1,1,\frac{D-2}{2}) \\ \hline \begin{array}{c}c^{a}\\ (1,1,\frac{D-2}{2})\end{array} & (1,1,\frac{D-2}{2}) & (2,0,\frac{D-2}{2}) & (0,0,\frac{D-2}{2}) \\ \begin{array}{c}\bar{c}^{\tilde{a}}\\ (-1,1,\frac{D-2}{2})\end{array} & (-1,1,\frac{D-2}{2}) & (0,0,\frac{D-2}{2}) & (-2,0,\frac{D-2}{2}) \end{array} \tag{120}$$

As first noted in Refs. [299,300], we see from the above that the degrees of freedom, ghost number and parity inherited by the products are very suggestive that squaring two BRST-covariant Yang–Mills theories results in the states, physical as well as first- and second-level ghosts, of a graviton, two-form and dilaton. Let us now make this precise at the linear level using the convolutive product [137,141]. For simplicity we adopt Feynman-'t Hooft gauge for both Yang–Mills factors and eliminate the Nakanishi-Lautrup fields, $b, \tilde{b}$, through their equations of motion. This is not required; arbitrary and independent gauge choices can be made [302]. The simplest[26] ansatz for the "gravity = gauge × gauge" dictionary is given by:

1. The graviton
$$h_{\mu\nu}=A_{(\mu}\circ\tilde{A}_{\nu)}-a\eta_{\mu\nu}(A^{\rho}\circ\tilde{A}_{\rho}-c^{\alpha}\circ\tilde{c}_{\alpha}). \tag{121}$$
2. The KR two-form
$$B_{\mu\nu}=A_{[\mu}\circ\tilde{A}_{\nu]}. \tag{122}$$
3. The dilaton
$$\varphi=A^{\rho}\circ\tilde{A}_{\rho}-c^{\alpha}\circ\tilde{c}_{\alpha}. \tag{123}$$

Note, here we have rescaled by A_{μ}, c^{α} (and similar on the right) by g^{-1} to ensure that the mass dimensions are consistent, that is $h_{\mu\nu}=A^{a}_{(\mu}\cdot\Phi_{a\tilde{a}}\cdot\tilde{A}^{\tilde{a}}_{\nu)}+\cdots$ and similar for the remaining fields. Note, a is left as a free parameter for now and the $\mathrm{SL}(2,\mathbb{R})$ ghost-antighost singlet $c^{\alpha}\circ\tilde{c}_{\alpha}=c^{\alpha}\circ\tilde{c}^{\beta}\varepsilon_{\alpha\beta}$ provides the trace of the graviton, but also contributes to the dilaton in Einstein frame.

Similarly, we have the ghost and ghost-for-ghost dictionaries:

[25] This follows from the requirement that the spectator is the convolutive pseudo-inverse of the ϕ^3 field: $\phi\cdot\Phi\cdot\phi=\phi$ implies $[\Phi]=(3D+2)/2$.

[26] Through the equations of motion this implies a mildly non-local relationship between the gauge and gravity sources [302]. Turning this around, imposing a simple gauge/gravity source dictionary, the equations of motion force non-local terms into the field dictionary [141].

1. Diffeomorphism (anti)ghost

$$c^\alpha_\mu = \frac{1}{2}\left(C^\alpha_\mu + \tilde{C}^\alpha_\mu\right) = \frac{1}{2}\left(c^\alpha \circ \tilde{A}_\mu + A_\mu \circ \tilde{c}^\alpha\right) \tag{124}$$

2. Two-form gauge (anti)ghost

$$\lambda^\alpha_\mu = \frac{1}{2}\left(C^\alpha_\mu - \tilde{C}^\alpha_\mu\right) = \frac{1}{2}\left(c^\alpha \circ \tilde{A}_\mu - A_\mu \circ \tilde{c}^\alpha\right). \tag{125}$$

3. Two-form gauge (anti)ghost-for-(anti)ghost

$$\lambda^{\alpha\beta} = c^{(\alpha} \circ \tilde{c}^{\beta)} = \begin{pmatrix} \lambda & \eta \\ \eta & \bar{\lambda} \end{pmatrix} \tag{126}$$

Having proposed a dictionary between the $\mathcal{N} = 0$ supergravity and pure Yang–Mills fields, the first consistency check is the BRST transformations. Said another way, the local gauge symmetries of the Yang–Mills factors must consistently generate the local diffeomorphism and 2-form gauge symmetries of h, B. Since we are working at linear level the non-Abelian Yang–Mills gauge group G breaks to $\dim \mathfrak{g}$ local $\mathrm{Un}(1)$ gauge symmetries and a global group, $G_{\text{global}} \cong G$,

$$\delta_{\epsilon,X} A = \epsilon dc + [A, X], \tag{127}$$

where ϵ, $\varepsilon(\epsilon) = 1$, $\mathrm{gh}(\epsilon) = -1$ is a constant parameter, $\delta_\epsilon = \epsilon Q$ and $X \in \mathfrak{g}_{\text{global}}$. Similarly, on the right factor $\tilde{G} \to \mathrm{Un}(1)^{\dim \tilde{\mathfrak{g}}} \times \tilde{G}_{\text{global}}$.

First, the gravity fields must be invariant under $G_{\text{global}} \times \tilde{G}_{\text{global}}$. This is trivially ensured by the spectator field, which transforms as

$$\delta_{X,\tilde{X}} \Phi = [\Phi, X] + [\Phi, \tilde{X}] \tag{128}$$

so that for any f, $\tilde{g}$ such that $\delta_X f = [f, X]$ and $\delta_{\tilde{X}} \tilde{g} = [\tilde{g}, \tilde{X}]$,

$$\delta_{X,\tilde{X}} f \circ \tilde{g} = 0, \tag{129}$$

which follows from the Killing form property $\langle X, [Y, Z]\rangle = \langle [X, Y], Z\rangle$.

Let us now turn to the BRST transformations: the linearised diffeomorphisms and the Abelian 2-form gauge transformations of $\mathcal{N} = 0$ supergravity. For convenience we recall here the linear BRST transformations in this gauge choice

$$QA = dc, \quad Qc = 0, \quad Q\bar{c} = -\partial A \tag{130}$$

and similar for the right factor. Then

$$\begin{aligned} \mathcal{Q}\varphi &= QA^\rho \circ \tilde{A}_\rho - Qc^\alpha \circ \tilde{c}_\alpha + A^\rho \circ \tilde{Q}\tilde{A}_\rho + c^\alpha \circ \tilde{Q}\tilde{c}_\alpha \\ &= \partial^\rho(c \circ \tilde{A}_\rho) - \partial^\rho(A_\rho \circ \tilde{c}) + \partial^\rho(A_\rho \circ \tilde{c}) - \partial^\rho(c \circ \tilde{A}_\rho) = 0 \end{aligned} \tag{131}$$

and

$$\begin{aligned}\mathcal{Q}(A_\mu \circ \tilde{A}_\nu) &= QA_\mu \circ \tilde{A}_\nu + A_\mu \circ \tilde{Q}\tilde{A}_\nu \\ &= \partial_\mu (c \circ \tilde{A}_\nu) + \partial_\nu (A_\mu \circ \tilde{c}) \\ &= \partial_\mu C_\nu + \partial_\nu \tilde{C}_\mu. \end{aligned} \tag{132}$$

Hence, from (125) we recover the linearised diffeomorphisms and 2-form gauge transformations:

$$\mathcal{Q}h_{\mu\nu} = \partial_\mu c_\nu + \partial_\nu c_\mu, \tag{133a}$$
$$\mathcal{Q}B_{\mu\nu} = \partial_\mu \lambda_\nu - \partial_\nu \lambda_\mu, \tag{133b}$$
$$\mathcal{Q}\varphi = 0. \tag{133c}$$

We see that the linearised general coordinate transformations of h, φ and the 2-form gauge symmetry of B are precisely recovered. Varying the ghost fields we obtain $\mathcal{Q}c_\mu = 0$ for the diffeomorphism ghost, as expected, and

$$\mathcal{Q}\lambda_\mu = \partial_\mu \lambda, \tag{134a}$$
$$\mathcal{Q}\lambda = 0, \tag{134b}$$
$$\mathcal{Q}\eta = \partial^\mu \lambda_\mu. \tag{134c}$$

These are precisely the gauge-for-gauge transformations of an Abelian 2-form [141, 303].[27] Note, this result relies on the Grassmann grading, strongly suggesting that the inclusion of ghosts is a necessary ingrediant. We are only left with the antighost transformations. These play the crucial role of mapping the gauging fixing choice of the Yang–Mills factors into the gravity theory. Recall, $Q\bar{c} = b$ and the equation of motion of b is determined by the gauge-fixing term. Similarly, focussing on the graviton, $\mathcal{Q}\bar{c}_\mu = b_\mu$, where b_μ is the 1-form Lagrange multiplier of the gauge-fixing action for the linearised Einstein–Hilbert action. The variation $\mathcal{Q}\bar{c}_\mu$ is determined by the antighost dictionary, (125), so that for our left/right Yang–Mills gauge choices with $a = 1/(D-2)$ we have,

$$\begin{aligned}\mathcal{Q}\bar{c}_\mu &= \frac{1}{2}\left(QA_\mu \circ \bar{\tilde{c}} + A_\mu \circ \tilde{Q}\bar{\tilde{c}} + Q\bar{c} \circ \tilde{A}_\mu - \bar{c} \circ \tilde{Q}\tilde{A}_\mu\right) \\ &= \frac{1}{2}\left(\partial_\mu (c^\alpha \circ \tilde{c}_\alpha) + \partial^\rho (A_\mu \circ \tilde{A}_\rho + A_\rho \circ \tilde{A}_\mu)\right) \\ &= -\partial^\rho h_{\rho\mu} + \frac{1}{2}\partial_\mu h, \end{aligned} \tag{135}$$

which corresponds to the de Donder linear diffeomorphism gauge-fixing function,

$$b^{(h)}_\mu = -\partial^\rho h_{\rho\mu} + \frac{1}{2}\partial_\mu h. \tag{136}$$

[27] The quantisation of the KB 2-form requires the full machinery of the BV formalism. See [156,157,303] for detailed accounts, the latter in the context of "gravity = gauge × gauge".

The requirement that $b_\mu^{(h)}$ is expressible in terms of gravity fields implies the de Donder term. The choice of $a = 1/(D-2)$ in (121) is fixed by the requirement that the diffeomorphism gauge-fixing is independent of the dilaton, reflecting our choice of Einstein frame. For arbitrary a there is a $\partial_\mu\varphi$ contribution to (136), while the de Donder term is left invariant. This is clearly a consequence of our choice of Yang–Mills gauge-fixing functions and the restriction to a local field dictionary. The point is that whatever choices we make for the left/right Yang–Mills theories they consistently map into the gravity theory; the gauge-fixing function of the gravity theory is determined by those of the Yang–Mills factors [302,304]. Similarly, the 2-form gauge-fixing term is given by

$$b_\mu^{(B)} = \partial_\mu\eta - \partial^\rho B_{\mu\rho}, \tag{137}$$

which is precisely the canonical gauge-fixing term of the Abelian 2-form [156,157, 303]. Equipped with the gauge-fixing terms we can then impose the equations of motion (including the gauge fixing) to uniquely determine the relationship between the Yang–Mills and $\mathcal{N} = 0$ supergravity sources, completing the linear dictionary [302,304].

Let us now work through the simplest example exhibiting all the local symmetries of interest, including supersymmetry. We consider the product of a left $D = 4, \mathcal{N} = 1$ super Yang–Mills multiplet and $D = 4$ Yang–Mills on the right. In this case we have the luxury of a full off-shell vector superfield for the left $\mathcal{N} = 1$ super Yang–Mills,

$$\begin{aligned} V(x,\theta,\bar\theta) =& M + i\theta\chi - i\bar\theta\bar\chi + i\theta^2 F - i\bar\theta^2\bar F - \theta\sigma^\mu\bar\theta A_\mu \\ &+ i\theta^2\bar\theta\left(\bar\psi + \frac{i}{2}\bar\sigma^\rho\partial_\rho\chi\right) - i\bar\theta^2\theta\left(\psi + \frac{i}{2}\sigma^\rho\partial_\rho\bar\chi\right) \\ &+ \frac{1}{2}\bar\theta^2\theta^2\left(D + \frac{1}{2}\Box M\right) \end{aligned} \tag{138}$$

transforming under local supergauge, non-Abelian global G and global super-Poincaré:

$$\delta V = \underbrace{C + \bar C}_{\text{local Abelian supergauge}} + \overbrace{[V, X]}^{\text{global non-Abelian } G} + \underbrace{\delta_\epsilon V}_{\text{global supersymmetry}} \tag{139}$$

where $C(x,\theta,\bar\theta)$ is a chiral superfield of ghosts

$$C(x,\theta,\bar\theta) = B + \sqrt{2}\theta\zeta + \theta^2 K + i\theta\sigma^\rho\bar\theta\partial_\rho c + \frac{i}{\sqrt{2}}\theta^2\bar\theta\bar\sigma^\rho\partial_\rho\zeta + \frac{1}{4}\theta^2\bar\theta^2\Box B \tag{140}$$

The product of $\mathcal{N} = 1$ with $\mathcal{N} = 0$ Yang–Mills generates $\mathcal{N} = 1$ supergravity coupled to a single chiral multiplet, as can be seen directly from the product of helicity states,

$$\left(+1, +\tfrac{1}{2}, -\tfrac{1}{2}, -1\right) \otimes (+1, -1) = \underbrace{\left(+2, +\tfrac{3}{2}, -\tfrac{3}{2}, -2\right)}_{\mathcal{N}=1 \text{ supergravity}} \oplus \underbrace{\left(+\tfrac{1}{2}, 0, 0, -\tfrac{1}{2}\right)}_{\mathcal{N}=1 \text{ supergravity}}. \tag{141}$$

The field and ghost dictionary is directly analogous:

$$\begin{aligned} H_\nu &= V \circ \tilde{A}_\nu \quad \text{real supergravity superfield} \\ S &= V \circ \tilde{c} \quad \text{real ghost superfield} \\ S_\nu &= C \circ \tilde{A}_\nu \quad \text{chiral ghost superfield} \end{aligned} \tag{142}$$

Varying the gravitational superfield via the dictionary

$$\delta H_\nu = S_\nu + \bar{S}_\nu + \partial_\nu S + \delta_\epsilon H_\nu. \tag{143}$$

This is the complete set of transformation rules for the new-minimal superfield at linearised approximation [305,306]. Hence, the local gravitational symmetries of general covariance, 2-form gauge invariance, local supersymmetry and local chiral symmetry follow from those of Yang–Mills at linear level. In particular, the product of the left fermion χ with the right ghost $\tilde{c}$ gives a local supergauge transformation of the gravitino. Schematically,

$$\psi_\nu = \psi \circ \tilde{A}_\nu, \quad \eta = \psi \circ \tilde{c} \quad \Rightarrow \quad \mathcal{Q}\psi_\nu = \partial_\mu \eta, \tag{144}$$

so that the presence of adjoint fermions induces local supersymmetries, in agreement with the BCJ double-copy. Including $\mathcal{N}$ and $\tilde{\mathcal{N}}$ adjoint fermions in the left and right factors, we obtain $\mathcal{N} + \tilde{\mathcal{N}}$ gravitini and local supergauge transformations and, hence, an $(\mathcal{N} + \tilde{N})$-extended supergravity theory.

The $12 + 12$ new minimal multiplet splits with respect to superconformal transformations into an $8 + 8$ conformal supergravity multiplet plus a $4 + 4$ conformal tensor multiplet,

$$\underbrace{\begin{pmatrix} \mathbf{5+3+1+3} \\ \mathbf{4+2+4+2} \end{pmatrix}}_{\text{new-minimal}} \to \underbrace{\begin{pmatrix} \mathbf{5+3} \\ \mathbf{4+4} \end{pmatrix}}_{\text{conformal}} + \underbrace{\begin{pmatrix} \mathbf{3+1} \\ \mathbf{2+2} \end{pmatrix}}_{\text{tensor}} \tag{145}$$

in terms of Spin(3) representations. Since the left (anti)ghost is a chiral superfield the ghost-antighost sector gives a compensating $4+4$ chiral (dilaton) multiplet [299,300], yielding old-minimal $12 + 12$ supergravity [307,308] coupled to a tensor multiplet, which, with the conventional 2-derivative Lagrangian, correctly corresponds to the on-shell content obtained by tensoring left/right helicity states.

This linear mapping can be used to construct a higher-order perturbative relationship using a BCJ-type formalism [302], building on that of [117]. Even in the absence of a full perturbative framework, the dictionary can be used to construct, for example, supersymmetric (single and multi-centre) black hole solutions in $\mathcal{N} = 2$ supergravity [115,116], in the weak-field limit. Finally, it can be applied to curved back-grounds, at least where the convolution can be made tractable [304].

3.2.2 Global symmetries

When coupled to other fields the Yang–Mills factors may have further global symmetries. In particular, in addition to global supersymmetry, pure $\mathcal{N}$-extended Yang–Mills

Table 4 U-dualities (global symmetries) of M-theory ($D = 11, \mathcal{N} = 1$ supergravity) compactified on an n-torus

n-torus	U-duality	$\mathcal{G}$	$\mathcal{H}$
1	$SO(1, 1, \mathbb{Z})$	$SO(1, 1, \mathbb{R})$	–
2	$SL(2, \mathbb{Z}) \times SO(1, 1, \mathbb{Z})$	$SL(2, \mathbb{R}) \times SO(1, 1, \mathbb{R})$	$SO(2, \mathbb{R})$
3	$SL(2, \mathbb{Z}) \times SL(3, \mathbb{Z})$	$SL(2, \mathbb{R}) \times SL(3, \mathbb{R})$	$SO(2, \mathbb{R}) \times SO(3, \mathbb{R})$
4	$SL(5, \mathbb{Z})$	$SL(5, \mathbb{R})$	$SO(5, \mathbb{R})$
5	$SO(5, 5, \mathbb{Z})$	$SO(5, 5, \mathbb{R})$	$SO(5, \mathbb{R}) \times SO(5, \mathbb{R})$
6	$E_{6(6)}(\mathbb{Z})$	$E_{6(6)}(\mathbb{R})$	$USp(8)$
7	$E_{7(7)}(\mathbb{Z})$	$E_{7(7)}(\mathbb{R})$	$SU(8)$
8	$E_{8(8)}(\mathbb{Z})$	$E_{8(8)}(\mathbb{R})$	$SO(16, \mathbb{R})$

theories always possess global R-symmetries. The global of the factors generate global symmetries in the product. In the context of supergravity, these typically take the form of non-compact global symmetries, $\mathcal{G}$, acting non-linearly on the scalar fields, the preeminent example being $D = 4, \mathcal{N} = 8$ supergravity, which has global symmetry $E_{7(7)}$, the maximally split non-compact real form of the second largest exceptional Lie group [309]. From the "open $\times$ open $=$ closed" string point of view, when it applies in the sense that the supergravity theory is the low energy effective field theory limit, these global symmetries can be understood as the continuous limit of the U-duality groups of M-theory [310]. This is indeed the case for all supergravity theories obtained from pure "$\mathcal{N}$ Yang–Mills $\times \tilde{\mathcal{N}}$ Yang–Mills" in any dimension, since the factors are all (possibly consistent truncations of) open string theories. In these examples, the scalar fields of the corresponding supergravity theory always parametrise a symmetric space $\mathcal{G}/\mathcal{H}$, where $\mathcal{H}$ is the maximal compact subgroup of the non-compact global symmetry group $\mathcal{G}$ [140]. This reflects and generalises our earlier observation that in $D = 4$ the axion–dilaton of $\mathcal{N} = 0$ supergravity belongs to $SL(2, \mathbb{R})/SO(2)$. An obvious question at this point is the Yang–Mills origin of such symmetries. In the following we shall describe this situation, revealing some unexpected surprises along the way, as well as some general principles both in the supersymmetric and non-supersymmetric cases. The question of global symmetries from squaring Yang–Mills has been addressed in, for example, [66,68,69,72,75,77,81,85,136,138–140,144,311,312], both in the context of scattering amplitudes and field theory.

As an example, in Table 4 we give the U-dualities of $D = 11$ M-theory compactified on an n-torus (equivalently $D = 10$ type IIA tring theory on an $(n - 1)$-torus), and the $\mathcal{G}, \mathcal{H}$ of the corresponding supergravity low energy effective field theory limits. These are the square of the maximally supersymmetric Yang–Mills thoeries in $D = 10 - (n + 1)$. As one observes, the global symmetries become increasingly manifest as one descends in dimension.[28] Thus, to fully expose the structure of the global symmetries with respect to squaring we should consider the product of super Yang–

[28] We stop at $D = 3$, which has $E_{8(8)}$ U-duality, the largest finite dimensional exceptional Lie algebra. One can continue to $D = 2, 1, 0$, invoking the infinite dimensional extended algebras $E_{9(9)}, E_{10(10)}, E_{11(11)}$ [313–316]. Although we will not discuss theses cases here, it would be interesting to investigate if they can be understood from the perspective of Yang–Mills squared.

Mills theories in $D = 3$. This was done in [136]. The result reveals a rather intriguing mathematical structure. The symmetry algebras obtained make up the Freudenthal–Rosenfeld–Tits magic square [317–319] as given in Table 6. As we shall explain this surprise has an elegant explanation, but first let us briefly return to the familiar case of $D = 4$ to make some generic observations.

Global symmetries: a first look To discuss global symmetries we can put aside the gauge/BRST transformations and focus on the asymptotic states of the on-shell spectrum. For the left/right Yang–Mills factors, these are labelled by their representations under the common spacetime little group $\mathrm{Spin}(D-2)$ and any internal global symmetries they may carry, which may include both R-symmetries R, $\tilde{R}$ and flavour groups F, $\tilde{F}$. We shall work at the infitesimal level, denoting the spacetime and internal Lie algebras by $\mathfrak{so}(D-2)$ and $\mathfrak{int}$, $\tilde{\mathfrak{int}}$. Suppressing the momentum label the states are denoted:

$$\text{left states} \quad |\rho_{\mathfrak{so}(D-2)};\, \rho_{\mathfrak{int}}\rangle_L^{\rho_{\mathfrak{g}}}, \quad \text{right states} \quad |\rho_{\mathfrak{so}(D-2)};\, \rho_{\tilde{\mathfrak{int}}}\rangle_L^{\rho_{\tilde{\mathfrak{g}}}}. \tag{146}$$

We have supressed the momentum and colour labels. However, it is sometimes important to recall the gauge group representation carried by the states, as indicated by the superscript. When all states are in the adjoint we will leave this implicit. Since the left/right spacetimes are identified, but the internal symmetries are not, the product states are $(\mathfrak{so}(D-2) \oplus \mathfrak{int} \oplus \tilde{\mathfrak{int}})$-modules.[29] Since, at tree-level all amplitudes of the left/right factors are invariant under $\mathfrak{int}$ and $\tilde{\mathfrak{int}}$, respectively, the global internal symmetry of the gravity theory is *at least* $\mathfrak{int} \oplus \tilde{\mathfrak{int}}$. In the absence of anomalies this persists to all orders in perturbation theory. But this need not be all symmetries of the gravitational theory. As we have argued, and will make explicit in the following, the global supersymmetries of the left ($\mathcal{N}$-extended) and right ($\tilde{\mathcal{N}}$-extended) factors sum to give local supersymmetries, so that the product theory has $(\mathcal{N}+\tilde{\mathcal{N}})$-extended local supersymmetry. Such theories have, at least, a linearly global symmetry isomorphic to the $(\mathcal{N}+\tilde{\mathcal{N}})$-extended R-symmetry group, which includes as a subgroup the product of the left and right R-symmetry groups. The product theory can and will (typically) have more symmetry than is present in its factors. From the perspective of the double-copy this is quite remarkable. The gravity amplitudes are built from the numerators of the factor only, which individually manifest only their own global symmetries, yet they conspire to yield the larger symmetry of the gravity theory.

Given enough symmetry in the factors, the symmetries of the product theory can be deduced unambiguously form the field theory product. First one must determine the linearised local symmetry as described in the previous section. Given this structure, one can then focuss on global symmetries in terms of either the fields or, more simply, the asymptotic states. Let us work through the paradigmatic example of "$\mathcal{N} = 4$ Yang–Mills $\times \tilde{\mathcal{N}} = 4$" Yang–Mills in $D = 4$. The $\mathcal{N} = 4$ Yang–Mills multiplet includes a gluon, four gluini and six scalars. The only allowed internal symmetry is the $\mathfrak{su}(4)$

[29] It might be elucidating to *not* identify the spacetime symmetries. For example, one can reformulate the $\mathcal{N} = 0$ supergravity action so as to have manifest left and right Lorentz symmetries [281,284]. Nonetheless, from the point of view of spectrum matching and global symmetries we should identify the spacetime little groups.

R-symmetry. The states are given by,

$$\begin{aligned}
\wedge^0 Q|1;\mathbf{1}\rangle &= |1;\mathbf{1}\rangle\\
\wedge^1 Q|1;\mathbf{1}\rangle &= |\tfrac{1}{2};\mathbf{4}\rangle\\
\wedge^2 Q|1;\mathbf{1}\rangle &= |0;\mathbf{6}\rangle\\
\wedge^3 Q|1;\mathbf{1}\rangle &= |-\tfrac{1}{2};\bar{\mathbf{4}}\rangle\\
\wedge^4 Q|1;\mathbf{1}\rangle &= |-1;\mathbf{1}\rangle
\end{aligned} \tag{147}$$

where $|h;\mathbf{n}\rangle$ denotes a helicity h state carrying $\mathfrak{su}(4)$ representation $\mathbf{n}$. We have also indicated the action of the supersymmetry charge $Q \sim |-\frac{1}{2};\mathbf{4}\rangle$. The product yields

	$\lvert 1;\mathbf{1}\rangle$	$\lvert \frac{1}{2};\mathbf{4}\rangle$	$\lvert 0;\mathbf{6}\rangle$	$\lvert -\frac{1}{2};\bar{\mathbf{4}}\rangle$	$\lvert -1;\mathbf{1}\rangle$
$\lvert 1;\mathbf{1}\rangle$	$\lvert 2;\mathbf{1},\mathbf{1}\rangle$	$\lvert \frac{3}{2};\mathbf{1},\mathbf{4}\rangle$	$\lvert 1;\mathbf{1},\mathbf{6}\rangle$	$\lvert \frac{1}{2};\mathbf{1},\bar{\mathbf{4}}\rangle$	$\lvert 0;\mathbf{1},\mathbf{1}\rangle$
$\lvert \frac{1}{2};\mathbf{4}\rangle$	$\lvert \frac{3}{2};\mathbf{4},\mathbf{1}\rangle$	$\lvert 1;\mathbf{4},\mathbf{4}\rangle$	$\lvert \frac{1}{2};\mathbf{4},\mathbf{6}\rangle$	$\lvert 0;\mathbf{4},\bar{\mathbf{4}}\rangle$	$\lvert -\frac{1}{2};\mathbf{4},\mathbf{1}\rangle$
$\lvert 0;\mathbf{6}\rangle$	$\lvert 1;\mathbf{6},\mathbf{1}\rangle$	$\lvert \frac{1}{2};\mathbf{6},\mathbf{4}\rangle$	$\lvert 0;\mathbf{6},\mathbf{6}\rangle$	$\lvert -\frac{1}{2};\mathbf{6},\bar{\mathbf{4}}\rangle$	$\lvert -1;\mathbf{6},\mathbf{1}\rangle$
$\lvert -\frac{1}{2};\bar{\mathbf{4}}\rangle$	$\lvert \frac{1}{2};\bar{\mathbf{4}},\mathbf{1}\rangle$	$\lvert 0;\bar{\mathbf{4}},\mathbf{4}\rangle$	$\lvert -\frac{1}{2};\bar{\mathbf{4}},\mathbf{6}\rangle$	$\lvert -1;\bar{\mathbf{4}},\bar{\mathbf{4}}\rangle$	$\lvert -\frac{1}{2};\bar{\mathbf{4}},\mathbf{1}\rangle$
$\lvert -1;\mathbf{1}\rangle$	$\lvert 0;\mathbf{1},\mathbf{1}\rangle$	$\lvert -\frac{1}{2};\mathbf{1},\mathbf{4}\rangle$	$\lvert -1;\mathbf{1},\mathbf{6}\rangle$	$\lvert -\frac{3}{2};\mathbf{1},\bar{\mathbf{4}}\rangle$	$\lvert -2;\mathbf{1},\mathbf{1}\rangle$

(148)

Gathering the positive helicity states we find they carry the $\mathfrak{int} \oplus \widetilde{\mathfrak{int}} \cong \mathfrak{su}(4) \oplus \mathfrak{su}(4)$ representations given by

$$2 \text{ Graviton:} \quad (\mathbf{1},\mathbf{1})_0 \quad \leftarrow \quad \mathbf{1} \tag{149}$$

$$\tfrac{3}{2} \text{ Gravitini:} \quad (\mathbf{4},\mathbf{1})_{\frac{1}{2}} + (\mathbf{1},\mathbf{4})_{-\frac{1}{2}} \quad \leftarrow \quad \mathbf{8} \tag{150}$$

$$1 \text{ Vectors:} \quad (\mathbf{6},\mathbf{1})_1 + (\mathbf{1},\mathbf{6})_{-1} + (\mathbf{4},\mathbf{4})_0 \quad \leftarrow \quad \mathbf{28} \tag{151}$$

$$\tfrac{1}{2} \text{ Spinors:} \quad (\bar{\mathbf{4}},\mathbf{1})_{\frac{3}{2}} + (\mathbf{1},\bar{\mathbf{4}})_{-\frac{3}{2}} + (\mathbf{6},\mathbf{4})_{\frac{1}{2}} + (\mathbf{4},\mathbf{6})_{-\frac{1}{2}} \quad \leftarrow \quad \mathbf{56} \tag{152}$$

$$0 \text{ Scalars:} \quad (\mathbf{1},\mathbf{1})_2 + (\mathbf{1},\mathbf{1})_{-2} + (\bar{\mathbf{4}},\mathbf{4})_1 + (\mathbf{4},\bar{\mathbf{4}})_{-1} + (\mathbf{6},\mathbf{6})_0 \quad \leftarrow \quad \mathbf{70} \tag{153}$$

while the negative helicity states carry their complex conjugates. We have indicated in the subscripts an additional $\mathfrak{u}(1)$ charge q corresponding to the difference, rather than sum, of the left and right helicities, $q = \tilde{h} - h$, so that $\mathfrak{so}(2) \oplus \widetilde{\mathfrak{so}(2)} \cong \mathfrak{so}(2) \oplus \mathfrak{u}(1)$, introduced in [72]. We have also indicated the branching under $\mathfrak{su}(8) \supset \mathfrak{u}(1) \oplus \mathfrak{su}(4) \oplus \mathfrak{su}(4)$ of the positive helicity $\mathcal{N} = 8$ supergravity states carrying the corresponding representations under its linearly realised global symmetry $\mathcal{H} \cong \mathrm{SU}(8)$ (we have weighted the $\mathfrak{u}(1)$ by a factor of 2 relative to the standard conventions). As observed in [72], with the $\mathfrak{u}(1)$ charges included, the spectrum of "$\mathcal{N} = 4$ Yang–Mills $\times \tilde{\mathcal{N}} = 4$" is precisely that of $\mathcal{N} = 8$ supergravity under $\mathfrak{su}(8) \supset \mathfrak{u}(1) \oplus \mathfrak{su}(4) \oplus \mathfrak{su}(4)$:

$$\begin{aligned}
\wedge^0 \mathcal{Q}|2;\mathbf{1}\rangle &= |1;\mathbf{1}\rangle\\
\wedge^1 \mathcal{Q}|2;\mathbf{1}\rangle &= |\tfrac{3}{2};\mathbf{8}\rangle\\
\wedge^2 \mathcal{Q}|2;\mathbf{1}\rangle &= |1;\mathbf{28}\rangle
\end{aligned}$$

$$\begin{aligned}\wedge^3 \mathcal{Q}|2;\mathbf{1}\rangle &= |\tfrac{1}{2};\overline{\mathbf{56}}\rangle\\ \wedge^4 \mathcal{Q}|2;\mathbf{1}\rangle &= |0;\mathbf{70}\rangle\\ \wedge^6 \mathcal{Q}|2;\mathbf{1}\rangle &= |-\tfrac{1}{2};\overline{\mathbf{56}}\rangle\\ \wedge^7 \mathcal{Q}|2;\mathbf{1}\rangle &= |-1;\overline{\mathbf{28}}\rangle\\ \wedge^8 \mathcal{Q}|2;\mathbf{1}\rangle &= |-2;\mathbf{1}\rangle.\end{aligned} \tag{154}$$

This also makes it clear that the left and right supercharges generate the $\mathcal{N}=8$ supercharges $\mathcal{Q}=Q\oplus\tilde{Q}$; under the product the supersymmetries sum $\mathcal{N}\times\tilde{\mathcal{N}}\to\mathcal{N}+\tilde{\mathcal{N}}$.

Note, unlike $\mathfrak{su}(4)\oplus\mathfrak{su}(4)$, the additional $\mathfrak{u}(1)$ was not *a priori* required to be a symmetry of the product theory. Nonetheless, its presence for the product of pure (super) Yang–Mills theories may be anticipated from various points of view. This analysis can be repeated for any product of pure super Yang–Mills theories, including $\mathcal{N}=0$, in $D=4$ [138,140]. In these cases, the additional $\mathfrak{u}(1)$ is always required at the level of symmetries. In particular, it ensures that the scalar manifold is symmetric [85], just as in the case of axion–dilaton gravity (dualised $\mathcal{N}=0$ supergravity) derived from the square of pure Yang–Mills. All such theories may be consistently truncated to $\mathcal{N}=0$ supergravity, so from this point of view it is reasonable, although not unquestionable, to expect the presence of the $\mathfrak{u}(1)$ symmetry. It is in this sense that the product of Yang–Mills theories generically yields symmetric scalar manifolds. Moreover, it is crucially present in all the double-copy constructed amplitudes in all cases that have been tested, although there is no formal proof that it holds to all points. For $\mathcal{N}+\tilde{\mathcal{N}}>4$ this had to be the case as the supergravity theories are unique and have this symmetry. For $\mathcal{N}+\tilde{\mathcal{N}}\le 4$ it is also present at tree-level, but is anomalous. Rather satisfyingly, these anomalies can be traced back to the factors through the double-copy [72,82]. As we shall review, it can also be understood from the division algebraic perspective on spacetime and supersymmetry.

We have thus far left the generators of $\mathfrak{su}(8)$ not contained in $\mathfrak{u}(1)\oplus\mathfrak{su}(4)\oplus\mathfrak{su}(4)$ unaccounted for. Although these cannot be generated by left or right transformations alone, they may be deduced from their product. Since we are seeking bosonic symmetries we can consider the “product” of the left and right supercharges, $Q\otimes\tilde{Q}$ to give us a map of $\mathcal{N}=8$ states that preserves helicity [140]. Let us denote the formally modified (by, roughly speaking, $\Box^{-1}$[30]) supersymmetry charges introduced in [140], by Q_-^a, Q_a^+, where the $\pm$ charge raises/lowers the helicity by $\pm 1/2$ and the superscript (subscript) a is in the $\mathbf{4}$ ($\bar{\mathbf{4}}$) of $\mathfrak{su}(4)$. The helicity preserving operators $Q_-^a\otimes\tilde{Q}_a^+$ and $Q_a^+\otimes\tilde{Q}_-^a$ operators sit in the $(\mathbf{4},\bar{\mathbf{4}})_1+(\bar{\mathbf{4}},\mathbf{4})_{-1}$ of $\mathfrak{u}(1)\oplus\mathfrak{su}(4)\oplus\mathfrak{su}(4)$, which matches the decomposition under $\mathfrak{su}(8)\supset\mathfrak{u}(1)\oplus\mathfrak{su}(4)\oplus\mathfrak{su}(4)$,

$$\mathbf{63}=(\mathbf{15},\mathbf{1})_0\oplus(\mathbf{1},\mathbf{15})_0\oplus(\mathbf{1},\mathbf{1})_0+(\mathbf{4},\bar{\mathbf{4}})_1+(\bar{\mathbf{4}},\mathbf{4})_{-1}. \tag{155}$$

The action of the helicity preserving operators $Q\otimes\tilde{Q}$ on the $\mathcal{N}=8$ states gives precisely the required transformations as described in [140]. Computing the commutators

[30] It is tempting to spectulate that this modification is related to the spectator scalar of the field product.

through this action, the full Lie algebra of $\mathfrak{su}(8)$ is recovered. This generalises to all dimensions and degrees of supersymmetry [140].

But this is not the end of the story. As emphasised, the equations of motion of $\mathcal{N}=8$ supergravity have a non-linear realised non-compact global symmetry, $E_{7(7)}$. If we make the assumption that the supergravity scalars parametrise a Riemannian symmetric homogenous space $\mathcal{G}/\mathcal{H}$, then $T_p(\mathcal{G}/\mathcal{H}) \cong \mathfrak{p}$, where $\mathfrak{g} = \mathfrak{h} + \mathfrak{p}$ for $\mathfrak{g}$, $\mathfrak{h}$ the Lie algebras of $\mathcal{G}$, $\mathcal{H}$ respectively. Then the previously derived $\mathfrak{h}$ representation carried by the scalars is enough to fix $\mathcal{G}$. In the present example we have $\mathbf{133} = \mathbf{63} + \mathbf{70}$ under $\mathfrak{su}(8) \subset \mathfrak{e}_{7(7)}$. Since the scalars are indeed in the $\mathbf{70}$ of $\mathfrak{su}(8)$, we infer the $E_{7(7)}$ and then can check the consistency of the representations carried by the vectors. Justifying the assumption that the scalars will parametrise a symmetric space by considering all $\mathcal{N}=0$ truncations, this provides a relatively systematic approach to fixing the global symmetries. In fact, it has a very natural algebraic/geometric underpinning as we shall describe in the following section. Of course, one could argue that given the uniqueness of $\mathcal{N} > 4$ supergravity, we knew the answer all along, however this is against the spirit of "gravity $=$ gauge $\times$ gauge". Besides, it can be generalised to $\mathcal{N} \leq 4$ supergravity theories, with a very large class of matter couplings, where we lose uniqueness and can drop the symmetric[31] scalar manifold assumption [85]. Nonetheless, one would like to see the generators in $\mathfrak{p} \subset \mathfrak{g}$ arise directly in terms of products of operators belonging to the left and right factors. Let us reconsider the tensor product of the (modified) supersymmetry charges introduced above. The ± 1 helicity states transform irreducibly in the $\mathbf{56}$ of $\mathfrak{e}_{7(7)}$, which decomposes into the $\mathbf{28} + \overline{\mathbf{28}}$ under $\mathfrak{su}(8)$. The $\mathbf{70} \in \mathfrak{e}_{7(7)} \ominus \mathfrak{su}(8)$ exchanges the helicity states so we require operators in the

$$\mathbf{70} = (\mathbf{1},\mathbf{1})_2 + (\mathbf{1},\mathbf{1})_{-2} + (\bar{\mathbf{4}},\mathbf{4})_1 + (\mathbf{4},\bar{\mathbf{4}})_{-1} + (\mathbf{6},\mathbf{6})_0 \tag{156}$$

of $\mathfrak{u}(1) \oplus \mathfrak{su}(4) \oplus \mathfrak{su}(4) \subset \mathfrak{su}(8)$ carrying helicity charge ± 2 related by conjugation and self-duality:

$$\begin{array}{rll} Q^+_a Q^+_b \otimes \tilde{Q}^+_{\tilde{a}} \tilde{Q}^+_{\tilde{b}} & & (\mathbf{6},\mathbf{6})^2_0 \\ Q^+_a Q^+_b Q^+_c \otimes \tilde{Q}^+_{\tilde{a}} & & (\mathbf{4},\bar{\mathbf{4}})^2_{-1} \\ Q^+_a \otimes \tilde{Q}^+_{\tilde{a}} \tilde{Q}^+_{\tilde{b}} \tilde{Q}^+_{\tilde{c}} & & (\bar{\mathbf{4}},\mathbf{4})^2_1 \\ Q^+_a Q^+_b Q^+_c Q^+_d \otimes \mathbb{1} & & (\mathbf{1},\mathbf{1})^{-2}_{-2} \\ \mathbb{1} \otimes \tilde{Q}^+_{\tilde{a}} \tilde{Q}^+_{\tilde{b}} \tilde{Q}^+_{\tilde{c}} \tilde{Q}^+_{\tilde{d}} & & (\mathbf{1},\mathbf{1})^2_2 \end{array} \tag{157}$$

which gives the correct commutation relations acting on the vector states. This is just suggestive and the full picture is yet to be made clear at the level of field theory.

Returning to amplitudes and the BCJ double-copy we can be much more concrete, since they intrinsically carry non-linearities. Generically, the presence of a non-linear symmetry of the scalar Lagrangian, we have in mind here $\mathcal{G}/\mathcal{H}$ sigma model, manifests itself through low-energy theorems regarding soft limits of the scalar amplitudes [311, 312,320]. If the space of scalars is a homogenous manifold, they are the Goldstone

[31] We must retain the homogenous assumption. In fact, as far as we are aware there are as yet no examples of double-copy constructible supergravity theories without homogenous scalar manifolds, despite there ubiquity for $\mathcal{N} \leq 2$.

bosons of $\mathcal{H} \subset \mathcal{G}$ and so are derivatively coupled. Consequently, on sending the momentum of any external scalar in any amplitude to zero the amplitude itself will vanish. This, in contrast, is not the case for the Yang–Mills factors. In the present case, the SU(4) symmetry of $\mathcal{N} = 4$ Yang–Mills is linearly realised on the scalars and single-soft-scalar limits do not vanish, but rather give the small-mass limit of Coulomb branch amplitudes [321]. Building amplitudes involving scalar through the double-copy of $\mathcal{N} = 4$ and then testing the vanishing of single-soft-scalar limits establishes that the scalars of the double-copy theory belong to a homogeneous space, despite the fact that scalars of the factors do not. One can go further by carefully considering taking double-soft limits in different orders to extract the commutation relations of the coset manifold and in this way piece together the full non-global symmetry of the double-copy theory, as has been done in some detail the $E_{7(7)}$ of $\mathcal{N} = 8$ supergravity [312]. The same principles can be applied to any case where the scalars belong to a homogenous manifold, a particularly elegant example constructed through the double-copy in [77] is given by the magic $\mathcal{N} = 2$ supergravity theories [237,244,245], which have exceptional non-compact global symmetries belonging to (a particular non-compact real form of) the Freudenthal–Rosenfeld–Tits magic square [317–319,322,323]. In fact, (a different non-compact real form of) the magic square arises naturally from "gravity = gauge × gauge" in a completely different context. This forms the next part of the story of the global symmetries.

3.2.3 Magic pyramids of symmetries

Recall, the U-duality symmetries grow as we descend in spacetime dimension. See Table 4. Let us therefore consider the global symmetries of the product of all pure $\mathcal{N} = 1, 2, 4, 8$ Yang–Mills theories in $D = 3$. Applying the principles of the preceding sections a remarkable result follows. It was shown in [136] that the resulting $(\mathcal{N}+\tilde{\mathcal{N}})$-extended supergravity theories have global symmetries given precisely by (a particular non-compact real form of) the Freudenthal–Rosenfeld–Tits magic square, as summarised in Table 5.

The Freudenthal–Rosenfeld–Tits magic square [317–319,322,323] is a 4×4 array $\mathfrak{m}(\mathbb{A}, \tilde{\mathbb{A}})$ of semi-simple Lie algebras given by pairs of composition algebras $\mathbb{A}, \tilde{\mathbb{A}} = \mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$, as given in Table 6. The original magic square was for compact real forms, but there are various modifications that allow for a variety of real forms. The complete set of possibilities are given in [324]. The magic square given in Table 6 can be concisely summarised by the magic square formula [136],

$$\mathfrak{ms}(\mathbb{A}_{\mathcal{N}_L}, \mathbb{A}_{\mathcal{N}_R}) := \mathfrak{tri}(\mathbb{A}_{\mathcal{N}_L}) \oplus \mathfrak{tri}(\mathbb{A}_{\mathcal{N}_R}) + 3(\mathbb{A}_{\mathcal{N}_L} \otimes \mathbb{A}_{\mathcal{N}_R}), \tag{158}$$

which adapts the compact version given in [196]. The triality algebra of $\mathbb{A}$, denoted $\mathfrak{tri}(\mathbb{A})$, is related to the total on-shell global symmetries of the associated super Yang–Mills theory [190]. This rather surprising connection, relating the magic square of Lie algebras to the square of super Yang–Mills, can be attributed to the existence of a unified $\mathbb{A}_{\mathcal{N}} = \mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ description of $D = 3$, $\mathcal{N} = 1, 2, 4, 8$ super Yang–Mills theories.

Table 5 $(\mathcal{N}+\tilde{\mathcal{N}})$-extended $D=3$ supergravities obtained by the product of left/right super Yang–Mills multiplets with $\mathcal{N}, \tilde{\mathcal{N}} = 1, 2, 4, 8$

$\mathcal{N}\backslash\tilde{\mathcal{N}}$	1	2	4	8
1	$\mathcal{G}=\mathrm{SL}(2,\mathbb{R})$	$\mathrm{SU}(2,1)$	$\mathrm{USp}(4,2)$	$F_{4(-20)}$
	$\mathcal{H}=\mathrm{SO}(2)$	$\mathrm{SO}(3)\times\mathrm{SO}(2)$	$\mathrm{SO}(5)\times\mathrm{SO}(3)$	$\mathrm{SO}(9)$
2	$\mathrm{SU}(2,1)$	$\mathrm{SU}(2,1)^2$	$\mathrm{SU}(4,2)$	$E_{6(-14)}$
	$\mathrm{SO}(3)\times\mathrm{SO}(2)$	$\mathrm{SO}(3)^2\times\mathrm{SO}(2)^2$	$\mathrm{SO}(6)\times\mathrm{SO}(3)\times\mathrm{SO}(2)$	$\mathrm{SO}(10)\times\mathrm{SO}(2)$
4	$\mathrm{USp}(4,2)$	$\mathrm{SU}(4,2)$	$\mathrm{SO}(8,4)$	$E_{7(-5)}$
	$\mathrm{SO}(5)\times\mathrm{SO}(3)$	$\mathrm{SO}(6)\times\mathrm{SO}(3)\times\mathrm{SO}(2)$	$\mathrm{SO}(8)\times\mathrm{SO}(3)\times\mathrm{SO}(3)$	$\mathrm{SO}(12)\times\mathrm{SO}(3)$
8	$F_{4(-20)}$	$E_{6(-14)}$	$E_{7(-5)}$	$E_{8(8)}$
	$\mathrm{SO}(9)$	$\mathrm{SO}(10)\times\mathrm{SO}(2)$	$\mathrm{SO}(12)\times\mathrm{SO}(3)$	$\mathrm{SO}(16)$

The algebras of the corresponding U-duality groups $\mathcal{G}$ and their maximal compact subgroups $\mathcal{H}$ are given by the magic square of Freudenthal–Rosenfeld–Tits [317–319,322,323]

Table 6 The magic square with real form corresponding to the product of pure super Yang–Mills theories in $D=3$ spacetime dimensions

$\mathbb{A}\backslash\tilde{\mathbb{A}}$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{O}$
$\mathbb{R}$	$\mathfrak{sl}(2,\mathbb{R})$	$\mathfrak{su}(2,1)$	$\mathfrak{sp}(4,2)$	$\mathfrak{f}_{4(-20)}$
$\mathbb{C}$	$\mathfrak{su}(2,1)$	$\mathfrak{su}(2,1)\times\mathfrak{su}(2,1)$	$\mathfrak{su}(4,2)$	$\mathfrak{e}_{6(-14)}$
$\mathbb{H}$	$\mathfrak{sp}(4,2)$	$\mathfrak{su}(4,2)$	$\mathfrak{so}(8,4)$	$\mathfrak{e}_{7(-5)}$
$\mathbb{O}$	$\mathfrak{f}_{4(-20)}$	$\mathfrak{e}_{6(-14)}$	$\mathfrak{e}_{7(-5)}$	$\mathfrak{e}_{8(8)}$

This observation was subsequently generalised to $D=3,4,6$ and 10 dimensions [138,190] by incorporating the well-known relationship between the existence of minimal super Yang–Mills theories in $D=3,4,6,10$ and the existence of the four division algebras $\mathbb{R},\mathbb{C},\mathbb{H},\mathbb{O}$ [187,189,191,195]. From this perspective the $D=3$ magic square forms the base of a "magic pyramid" of supergravities, given in Fig. 1. The Lie algebras are given by the *magic pyramid formula*:

$$\mathfrak{mp}(\mathbb{A}_n,\mathbb{A}_{n\mathcal{N}},\mathbb{A}_{n\tilde{\mathcal{N}}}) := \left\{u\in\mathfrak{ms}(\mathbb{A}_{n\mathcal{N}},\mathbb{A}_{n\tilde{\mathcal{N}}})\ominus\mathfrak{so}(\mathbb{A}_n)_{ST}\,\middle|\,[u,\mathfrak{so}(\mathbb{A}_n)_{ST}]=0\right\}. \tag{159}$$

These constructions build on a long line of work relating division algebras and magic squares to spacetime and supersymmetry. See [35,187–189,191,195,237,244,245,313,324–358] for a glimpse of the relevant literature. An early example[32] in the context of group disintegrations in supergravity appears in [313]. Before developing these ideas we should take a breif detour through division algebras and the magic square.

Division algebras and the magic square In this section we follow closely [195,196]; we refer the reader to these works for more detailed explanations and proofs. An algebra $\mathbb{A}$ defined over $\mathbb{R}$ with identity element e_0, is said to be *composition* if it has

[32] As far as we are aware the first instance in this context.

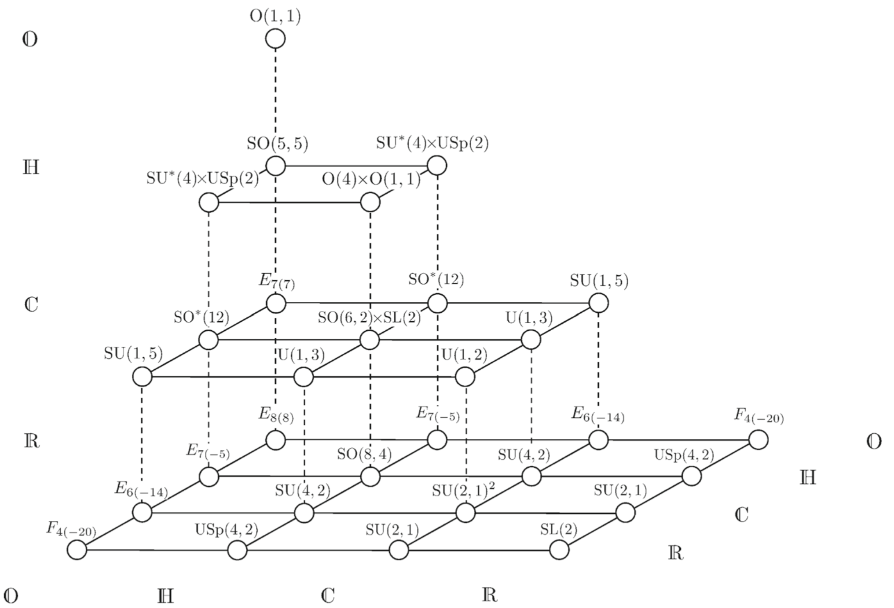


Fig. 1 A magic pyramid of supergravities. The vertical axis labels the spacetime division algebra $\mathbb{A}_n$, while the horizontal axes label the algebras associated with the number of supersymmetries $\mathbb{A}_{n\mathcal{N}}$ and $\mathbb{A}_{n\tilde{\mathcal{N}}}$

a non-degenerate quadratic form[33] $\mathbf{n} : \mathbb{A} \to \mathbb{R}$ such that,

$$\mathbf{n}(ab) = \mathbf{n}(a)\mathbf{n}(b), \quad \forall\ a, b \in \mathbb{A}, \tag{160}$$

where we denote the multiplicative product of the algebra by juxtaposition. Regarding $\mathbb{R} \subset \mathbb{A}$ as the scalar multiples of the identity $\mathbb{R}e_0$ we may decompose $\mathbb{A}$ into its "real" and "imaginary" parts $\mathbb{A} = \mathbb{R} \oplus \text{Im}\mathbb{A}$, where $\text{Im}\mathbb{A} \subset \mathbb{A}$ is the subspace orthogonal to $\mathbb{R}$. An arbitrary element $a \in \mathbb{A}$ may be written $a = \text{Re}(a) + \text{Im}(a)$. Here $\text{Re}(a) \in \mathbb{R}e_0$, $\text{Im}(a) \in \text{Im}\mathbb{A}$. Defining conjugation using the bilinear form,

$$\tau(a) \equiv \overline{a} := \langle a, e_0 \rangle e_0 - a, \quad \langle a, b \rangle := \mathbf{n}(a+b) - \mathbf{n}(a) - \mathbf{n}(b). \tag{161}$$

we as usual write

$$\text{Re}(a) = \frac{1}{2}(a + \overline{a}), \quad \text{Im}(a) = \frac{1}{2}(a - \overline{a}). \tag{162}$$

A composition algebra $\mathbb{A}$ is said to be *division* if it contains no zero divisors,

$$ab = 0 \quad \Rightarrow \quad a = 0 \quad \text{or} \quad b = 0,$$

in which case $\mathbf{n}$ is positive semi-definite and $\mathbb{A}$ is referred to as a normed division algebra. Hurwitz's celebrated theorem states that there are exactly four normed division

33 A *quadratic norm* on a vector space V over a field $\mathbb{R}$ is a map $\mathbf{n} : V \to \mathbb{R}$ such that: (1) $\mathbf{n}(\lambda a) = \lambda^2 \mathbf{n}(a)$, $\lambda \in \mathbb{R}$, $a \in V$ and (2) $\langle a, b \rangle := \mathbf{n}(a+b) - \mathbf{n}(a) - \mathbf{n}(b)$ is bilinear.

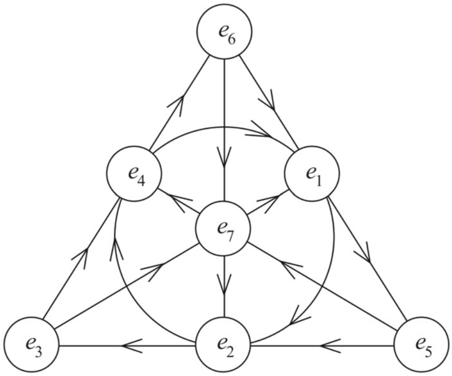


Fig. 2 The Fano plane. The structure constants are determined by the Fano plane, $C_{ijk} = 1$ if ijk lies on a line and is ordered according as its orientation. Each oriented line follows the rules of quaternionic multiplication. For example, $e_2e_3 = e_5$ and cyclic permutations; odd permutations go against the direction of the arrows on the Fano plane and we pick up a minus sign, e.g. $e_3e_2 = -e_5$

algebras [359]: the reals, complexes, quaternions and octonions, denoted respectively by $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ and $\mathbb{O}$. They may be constructed via the Cayley-Dickson doubling procedure, $\mathbb{A}' = \mathbb{A} \oplus \mathbb{A}$ with multiplication in $\mathbb{A}'$ defined by

$$(a, b)(c, d) = (ac - d\bar{b}, \bar{a}d + cb). \tag{163}$$

With each doubling a property is lost as summarised here:

	dim	Division	Associative	Commutative	Ordered
$\mathbb{R} = \mathbb{R}$	1	*yes*	*yes*	*yes*	*yes*
$\mathbb{C} \cong \mathbb{R} \oplus \mathbb{R}$	2	*yes*	*yes*	*yes*	*no*
$\mathbb{H} \cong \mathbb{C} \oplus \mathbb{C}$	4	*yes*	*yes*	*no*	*no*
$\mathbb{O} \cong \mathbb{H} \oplus \mathbb{H}$	8	*yes*	*no*	*no*	*no*

Note that, while the octonions are not associative they are alternative:

$$[a, b, c] := (ab)c - a(bc) \tag{164}$$

is an alternating function under the interchange of its arguments. This property is crucial for supersymmetry.

An element $a \in \mathbb{O}$ may be written $a = a^a e_a$, where $a = 0, \ldots, 7$, $a^a \in \mathbb{R}$ and $\{e_a\}$ is a basis with one real e_0 and seven e_i, $i = 1, \ldots, 7$, imaginary elements. The octonionic multiplication rule is,

$$e_a e_b = (\delta_{a0}\delta_{bc} + \delta_{0b}\delta_{ac} - \delta_{ab}\delta_{0c} + C_{abc})\, e_c, \tag{165}$$

where C_{abc} is totally antisymmetric and $C_{0bc} = 0$. The non-zero C_{ijk} are given by the Fano plane. See Fig. 2.

There are three symmetry algebras on $\mathbb{A}$ that we will make use of here. The *norm preserving* algebra is defined as,

$$\mathfrak{so}(\mathbb{A}) := \{A \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{A}) | \langle Aa, b\rangle + \langle a, Ab\rangle = 0,\ \forall a, b \in \mathbb{A}\}, \tag{166}$$

yielding,

$$
\begin{aligned}
\mathfrak{so}(\mathbb{R}) &\cong \emptyset, \\
\mathfrak{so}(\mathbb{C}) &\cong \mathfrak{so}(2), \\
\mathfrak{so}(\mathbb{H}) &\cong \mathfrak{so}(3) \oplus \mathfrak{so}(3), \\
\mathfrak{so}(\mathbb{O}) &\cong \mathfrak{so}(8).
\end{aligned} \tag{167}
$$

The *triality* algebra of $\mathbb{A}$ is defined as triples $(A, B, C) \in \mathfrak{so}(\mathbb{A}) \oplus \mathfrak{so}(\mathbb{A}) \oplus \mathfrak{so}(\mathbb{A})$ that act as generalise derivations,

$$
\mathfrak{tri}(\mathbb{A}) := \{(A, B, C) | A(ab) = B(a)b + aC(b), \ \forall a, b \in \mathbb{A}\}, \tag{168}
$$

yielding,

$$
\begin{aligned}
\mathfrak{tri}(\mathbb{R}) &\cong \emptyset, \\
\mathfrak{tri}(\mathbb{C}) &\cong \mathfrak{so}(2) \oplus \mathfrak{so}(2), \\
\mathfrak{tri}(\mathbb{H}) &\cong \mathfrak{so}(3) \oplus \mathfrak{so}(3) \oplus \mathfrak{so}(3), \\
\mathfrak{tri}(\mathbb{O}) &\cong \mathfrak{so}(8).
\end{aligned} \tag{169}
$$

Note, the octonionic triality algebra reduces to a single copy of $\mathfrak{so}(8)$. This is the statement of infinitesimal triality: for all $A \in \mathfrak{so}(8)$ there exist unique B, C such that

$$
A(ab) = B(a)b + aC(b), \quad \forall a, b \in \mathbb{O}. \tag{170}
$$

One can regard the triality algebra as a generalised form of the *derivation* algebra,

$$
\mathfrak{der}(\mathbb{A}) = \{A \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{A}) | A(ab) = A(a)b + aA(b)\}, \tag{171}
$$

which for $\mathbb{A} = \mathbb{O}$ gives the smallest exceptional Lie algebra,

$$
\begin{aligned}
\mathfrak{der}(\mathbb{R}) &\cong \emptyset, \\
\mathfrak{der}(\mathbb{C}) &\cong \emptyset, \\
\mathfrak{der}(\mathbb{H}) &\cong \mathfrak{so}(3), \\
\mathfrak{der}(\mathbb{O}) &\cong \mathfrak{g}_{2(-14)}.
\end{aligned} \tag{172}
$$

This provides the first example of a division algebraic description of an exceptional Lie algebra. In fact, the entire magic square can be realised in terms of the division algebras. The magic square was the result of an effort to give a unified and geometrically motivated description of Lie algebras, including the remaining exceptional cases of $\mathfrak{f}_4, \mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8$. The classical Lie algebras $\mathfrak{so}(n)$, $\mathfrak{su}(n)$, $\mathfrak{sp}(n)$ are very naturally captured by $\mathbb{R}, \mathbb{C}, \mathbb{H}$ geometrical structures, respectively. There are a number of ways of articulating this idea, but perhaps the most concise is in terms of the isometries of projective geometries. In particular, the isometry Lie algebras are:

$$\mathfrak{Isom}(\mathbb{RP}^n) \cong \mathfrak{so}(n+1), \quad \mathfrak{Isom}(\mathbb{CP}^n) \cong \mathfrak{su}(n+1), \quad \mathfrak{Isom}(\mathbb{HP}^n) \cong \mathfrak{sp}(n+1). \tag{173}$$

This sequence is rather suggestive; can we continue it to include $\mathfrak{Isom}(\mathbb{OP}^n)$? Despite non-associativity it was shown by Moufang [360] that one can consistently construct the octonionic projective line and plane, $\mathbb{OP}^1$ and $\mathbb{OP}^2$. The latter is often referred to as the Cayley plane. However, we cannot go beyond $n = 2$ for the octonions,[34] which in this context reflects the fact that there is indeed just a finite set of exceptional Lie algebras not belonging to any countably infinite family. The $\mathbb{OP}^1$ example is constructed in direct analogy with the real, complex and quaternionic cases.[35] It has $\mathfrak{Isom}(\mathbb{OP}^1) \cong \mathfrak{so}(8)$, so does not give us anything new. The octonionic plane has a more intricate structure. An element $(a, b, c) \in \mathbb{O}^3$ with $\mathbf{n}(a) + \mathbf{n}(a) + \mathbf{n}(c) = 1$ and $(ab)c = a(bc)$ gives a point in $\mathbb{OP}^2$, the line through the origin containing (a, b, c) in $\mathbb{O}^3$. It is not difficult to show the space of such elements is a 16-dimensional real manifold embedded in $\mathbb{O}^3$ through eight real constraints: $\mathbf{n}(a) + \mathbf{n}(a) + \mathbf{n}(c) = 1$ and $(ab)c = a(bc)$. The lines in $\mathbb{OP}^2$ are copies of $\mathbb{OP}^1$ and there is a duality relation sending lines/points into points/lines preserving the incidence structure. Borel showed that $F_{4(-52)}$ is the isometry group of a 16-dimensional projective plane, which is none other than $\mathbb{OP}^2$. One can show that the points and lines in $\mathbb{OP}^2$ are in one-to-one incidence preserving correspondence with trace 1 and 2 projectors in the Jordan algebra of 3×3 octonionic Hermitian matrices $\mathfrak{J}_3(\mathbb{O})$ (treating projectors as propositions the incidence relation in $\mathfrak{J}_3^{\mathbb{O}}$ is given by implication) [361]. Then $F_{4(-52)} = \text{Isom}(\mathbb{OP}^2)$ follows automatically from the result of Chevalley and Schafer that $F_{4(-52)} = \text{Aut}(\mathfrak{J}_3(\mathbb{O}))$, the group preserving the Jordan product with Lie algebra $\mathfrak{der}(\mathfrak{J}_3(\mathbb{O}))$ [362]. In summary, the sequence in (173) is continued to include,

$$\mathfrak{Isom}(\mathbb{OP}^2) \cong \mathfrak{der}(\mathfrak{J}_3(\mathbb{O})) \cong \mathfrak{f}_{4(-52)}. \tag{174}$$

Since $F_{4(-52)}$ acts transitively on the space of trace 1 projectors and the stabiliser of a given trace 1 projector is isomorphic to Spin(9) we have,

$$\mathbb{OP}^2 \cong F_{4(-52)}/\text{Spin}(9). \tag{175}$$

The Cayley plane is a homogenous symmetric space with $T_p(\mathbb{OP}^2) \cong \mathbb{O}^2$, which carries the spinor representation of Spin(9); under $F_{4(-52)} \supset \text{Spin}(9)$ we have

$$\begin{aligned} \mathfrak{f}_{4(-52)} &\cong \mathfrak{so}(\mathbb{R} \oplus \mathbb{O}) + \mathbb{O}^2 \\ &\cong \mathfrak{so}(\mathbb{O}) + \mathbb{O} + \mathbb{O} + \mathbb{O}. \end{aligned} \tag{176}$$

[34] One way to understand this is in terms of Jordan algebras. Points in $\mathbb{OP}^2$ are bijectively identified with trace 1 projectors in $\mathfrak{J}_3^{\mathbb{O}}$, the Jordan algebra of 3×3 octonionic Hermitian matrices. However, for $m > 3$, $m \times m$ octonionic Hermitian matrices do not form a Jordan algebra.

[35] Non-associativity, however, implies that the line through the origin containing the point (a, b) is not given by $\{(\alpha a, \alpha b) | \alpha \in \mathbb{O}\}$, unless $x = 1$ or $y = 1$. This obstacle is easily avoided as all non-zero octonions have an inverse; (a, b) is equivalent to $(b^{-1}a, 1)$ or $(1, a^{-1}b)$ for $b \neq 0$ or $a \neq 0$, giving two charts with a smooth transition function on their overlap. See [195].

The three $\mathbb{O}$ terms in the final line transform in the three triality related 8-dimensional representations of $\mathfrak{so}(8)$, the vector, spinor and conjugate spinor. It is this triality relation which implies that $\mathfrak{tri}(\mathbb{O}) \cong \mathfrak{so}(\mathbb{O})$.

Seemingly inspired by the trivial identity $\mathbb{O} \cong \mathbb{R} \otimes \mathbb{O}$ Boris Rosenfeld [319] proposed a natural extension of this construction,

$$\begin{aligned}
\mathfrak{Isom}((\mathbb{C} \otimes \mathbb{O})\mathbb{P}^2) &\cong \mathfrak{e}_{6(-78)}, \\
\mathfrak{Isom}((\mathbb{H} \otimes \mathbb{O})\mathbb{P}^2) &\cong \mathfrak{e}_{7(-133)}, \\
\mathfrak{Isom}((\mathbb{O} \otimes \mathbb{O})\mathbb{P}^2) &\cong \mathfrak{e}_{8(-248)},
\end{aligned} \tag{177}$$

thus giving a uniform geometric description for all Lie algebras. The would-be tangents spaces $(\mathbb{A} \otimes \mathbb{O})^2$ have the correct dimensions. However, it is not actually possible to construct projective spaces over $\mathbb{H} \otimes \mathbb{O}$ and $\mathbb{O} \otimes \mathbb{O}$ using the logic applied to $\mathbb{O}\mathbb{P}^2$, essentially because they do not yield Jordan algebras. They nonetheless can be identified with Riemannian geometries with isometries $E_{7(-133)}$ and $E_{8(-248)}$, respectively. Indeed, the Lie algebra decompositions,[36]

$$\begin{aligned}
\mathfrak{f}_{4(-52)} &\cong \mathfrak{so}(\mathbb{R} \oplus \mathbb{O}) + (\mathbb{R} \otimes \mathbb{O})^2 \\
\mathfrak{e}_{6(-78)} &\cong \mathfrak{so}(\mathbb{C} \oplus \mathbb{O}) \oplus \mathfrak{u}(1) + (\mathbb{C} \otimes \mathbb{O})^2 \\
\mathfrak{e}_{7(-133)} &\cong \mathfrak{so}(\mathbb{H} \oplus \mathbb{O}) \oplus \mathfrak{sp}(1) + (\mathbb{H} \otimes \mathbb{O})^2 \\
\mathfrak{e}_{8(-248)} &\cong \mathfrak{so}(\mathbb{O} \oplus \mathbb{O}) + (\mathbb{O} \otimes \mathbb{O})^2
\end{aligned} \tag{178}$$

naturally suggest the identifications

$$\begin{aligned}
\mathrm{Isom}((\mathbb{R} \otimes \mathbb{O})\mathbb{P}^2) &= F_{4(-52)}/\mathrm{Spin}(9) \\
\mathrm{Isom}((\mathbb{C} \otimes \mathbb{O})\mathbb{P}^2) &= E_{6(-78)}/[(\mathrm{Spin}(10) \times \mathrm{Un}(1))/\mathbb{Z}_4] \\
\mathrm{Isom}((\mathbb{H} \otimes \mathbb{O})\mathbb{P}^2) &= E_{7(-133)}/[(\mathrm{Spin}(10) \times \mathrm{Sp}(1))/\mathbb{Z}_2] \\
\mathrm{Isom}((\mathbb{O} \otimes \mathbb{O})\mathbb{P}^2) &= E_{8(-248)}/[\mathrm{Spin}(16)/\mathbb{Z}_2]
\end{aligned} \tag{179}$$

with tangent spaces $(\mathbb{R}\otimes\mathbb{O})^2$, $(\mathbb{C}\otimes\mathbb{O})^2$, $(\mathbb{H}\otimes\mathbb{O})^2$, $(\mathbb{O}\otimes\mathbb{O})^2$ carrying the appropriate spinor representations. Using the Tits' construction [318] the isometry algebras are given by the natural generalisation of (174),

$$\begin{aligned}
\mathfrak{f}_{4(-52)} &\cong \mathfrak{der}(\mathbb{R}) \oplus \mathfrak{der}(\mathfrak{J}_3(\mathbb{O})) + \mathrm{Im}\mathbb{R} \otimes \mathfrak{J}'_3(\mathbb{O}) \\
\mathfrak{e}_{6(-78)} &\cong \mathfrak{der}(\mathbb{C}) \oplus \mathfrak{der}(\mathfrak{J}_3(\mathbb{O})) + \mathrm{Im}\mathbb{C} \otimes \mathfrak{J}'_3(\mathbb{O}) \\
\mathfrak{e}_{7(-133)} &\cong \mathfrak{der}(\mathbb{H}) \oplus \mathfrak{der}(\mathfrak{J}_3(\mathbb{O})) + \mathrm{Im}\mathbb{H} \otimes \mathfrak{J}'_3(\mathbb{O}) \\
\mathfrak{e}_{8(-248)} &\cong \mathfrak{der}(\mathbb{O}) \oplus \mathfrak{der}(\mathfrak{J}_3(\mathbb{O})) + \mathrm{Im}\mathbb{O} \otimes \mathfrak{J}'_3(\mathbb{O}),
\end{aligned} \tag{180}$$

[36] Note, the additional factors are given by intermediate algebras: $\mathfrak{tri}(\mathbb{A}) \ominus \mathfrak{int}(\mathbb{A}) = \emptyset, \mathfrak{u}(1), \mathfrak{sp}(1), \emptyset$ for $\mathbb{A} = \mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ [196].

Table 7 The magic square given by the Tits' construction

$\otimes$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{O}$
$\mathbb{R}$	$\mathfrak{su}(2)$	$\mathfrak{su}(3)$	$\mathfrak{sp}(6)$	$\mathfrak{f}_{4(-52)}$
$\mathbb{C}$	$\mathfrak{su}(3)$	$\mathfrak{su}(3)\times\mathfrak{su}(3)$	$\mathfrak{su}(6)$	$\mathfrak{e}_{6(-78)}$
$\mathbb{H}$	$\mathfrak{sp}(6)$	$\mathfrak{su}(6)$	$\mathfrak{so}(12)$	$\mathfrak{e}_{7(-133)}$
$\mathbb{O}$	$\mathfrak{f}_{4(-52)}$	$\mathfrak{e}_{6(-78)}$	$\mathfrak{e}_{7(-133)}$	$\mathfrak{e}_{8(-248)}$

where $\mathfrak{J}'$ denotes the subset of traceless elements in $\mathfrak{J}$. Generalising further, the Tits' construction defines a Lie algebra,

$$\mathfrak{ms}(\mathbb{A},\tilde{\mathbb{A}})_{\text{compact}} := \mathfrak{der}(\mathbb{A})\oplus\mathfrak{der}(\mathfrak{J}_3(\tilde{\mathbb{A}})) + \text{Im}\mathbb{A}\otimes\mathfrak{J}'_3(\tilde{\mathbb{A}}), \tag{181}$$

which yields the compact magic square given in Table 7. The "magic" is that Table 7 symmetric about the diagonal despite the apparent asymmetry of (181). To obtain a magic square with the non-compact real forms that follow from squaring Yang–Mills, as given in Table 6, one can use a Lorentzian Jordan algebra [324],

$$\mathfrak{ms}(\mathbb{A},\tilde{\mathbb{A}}) := \mathfrak{der}(\mathbb{A})\oplus\mathfrak{der}(\mathfrak{J}_{1,2}(\tilde{\mathbb{A}})) + \text{Im}\mathbb{A}\otimes\mathfrak{J}'_{1,2}(\tilde{\mathbb{A}}). \tag{182}$$

The commutation relations are omitted here, as later we shall see that Yang–Mills squared gives an alternative form of (182), based on the Barton-Sudbery triality construction [196], that is manifestly symmetric in $\mathbb{A}$, $\tilde{\mathbb{A}}$ [136,138], for which we will present the details in full. This symmetric form reflects the fact that the squaring procedure is itself symmetric on interchanging the left and right theories.

Division algebras and Yang–Mills theories: In the two previous sections we saw that the "square" of $D = 3$ super Yang–Mills theories and the "square" of division algebras both led to the magic square of Freudenthal. Surely this is no coincidence. Indeed, there is a long history of work connecting supersymmetry, spacetime and the division algebras [35,187–189,191,195,237,244,245,324–339,341–343,346,348,352, 353,356–358], which, as we shall review, underlies this magical meeting.

Perhaps the most direct link from division algebras to spacetime symmetries comes via the Lie algebra isomorphism of Sudbery [191],

$$\mathfrak{sl}(2,\mathbb{A}) \cong \mathfrak{so}(1, 1+\dim\mathbb{A}), \tag{183}$$

which identifies $D = 3, 4, 6, 10$ as algebraically special. This is itself tied to the earlier observation of Kugo and Townsend [187] that the existence of minimal super Yang–Mills multiplets in only $D = 3, 4, 6, 10$ is related to the uniqueness of $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$. This was followed-up by a number of authors [192,193,363–365], sharpening the correspondence between supersymmetry and division algebras. The final case of $D = 10$, $\mathbb{A} = \mathbb{O}$ was developed most carefully in [189], where the link between supersymmetry and the alternativity of $\mathbb{O}$ was emphasised.

Pulling together these ideas, it was shown in [190] that $\mathcal{N}$-extended super Yang–Mills theories in $D = n + 2$ dimensions are completely specified (the field content,

Table 8 A table of algebras: $\mathfrak{sym}(\mathbb{A}_n, \mathbb{A}_{n\mathcal{N}})$

$\mathbb{A}_n \backslash \mathbb{A}_{n\mathcal{N}}$	$\mathbb{O}$	$\mathbb{H}$	$\mathbb{C}$	$\mathbb{R}$
$\mathbb{O}$	$\mathfrak{so}(8)_{ST}$			
$\mathbb{H}$	$\mathfrak{so}(4)_{ST} \oplus \mathfrak{sp}(1) \oplus \mathfrak{sp}(1)$	$\mathfrak{so}(4)_{ST} \oplus \mathfrak{sp}(1)$		
$\mathbb{C}$	$\mathfrak{so}(2)_{ST} \oplus \mathfrak{su}(4)$	$\mathfrak{so}(2)_{ST} \oplus \mathfrak{sp}(1) \oplus \mathfrak{so}(2)$	$\mathfrak{so}(2)_{ST} \oplus \mathfrak{so}(2)$	
$\mathbb{R}$	$\mathfrak{so}(8)$	$\mathfrak{so}(4) \oplus \mathfrak{sp}(1)$	$\mathfrak{so}(2) \oplus \mathfrak{so}(2)$	$\emptyset$

This lets us read off the spacetime and internal symmetries in each Yang–Mills theory. For example, one can see the familiar R-symmetries in $D = 4$: U(1), U(2) and SU(4) for $\mathcal{N} = 1, 2, 4$, respectively. Note that the symmetries in $D = 3$ are entirely internal and that they include the R-symmetry as a subgroup (these are actually the symmetries of the theories after dualising the vector to a scalar)

Lagrangian and transformation rules) by selecting an ordered pair of division algebras: $\mathbb{A}_n$ for the spacetime dimension and $\mathbb{A}_{n\mathcal{N}}$ for the degree of supersymmetry, where the subscripts denote the dimension of the algebras. Consequently, the dual appearances of the magic square in $D = 3$, or equivalently for $\mathbb{A}_n = \mathbb{R}$, can be explained by the observation that $D = 3, \mathcal{N} = 1, 2, 4, 8$ Yang–Mills theories can be formulated with a single Lagrangian and a single set of transformation rules, using fields valued in $\mathbb{R}, \mathbb{C}, \mathbb{H}$ and $\mathbb{O}$, respectively [136]. Tensoring an $\mathbb{A}$-valued $D = 3$ super Yang–Mills multiplet with an $\tilde{\mathbb{A}}$-valued $D = 3$ super Yang–Mills multiplet yields a $D = 3$ supergravity multiplet with fields valued in $\mathbb{A} \otimes \tilde{\mathbb{A}}$, making a magic square of U-dualities appear rather natural.

As noted in [190], the overall (spacetime little group plus internal) symmetry of the $\mathcal{N} = 1$ theory in $D = n + 2$ dimensions is given by the triality algebra, $\mathfrak{tri}(\mathbb{A}_n)$. If we dimensionally reduce these theories we obtain super Yang–Mills with $\mathcal{N}$ supersymmetries whose overall symmetries are given by,

$$\mathfrak{sym}(\mathbb{A}_n, \mathbb{A}_{n\mathcal{N}}) := \left\{(A, B, C) \in \mathfrak{tri}(\mathbb{A}_{n\mathcal{N}}) | [A, \mathfrak{so}(\mathbb{A}_n)_{ST}] = 0, \quad \forall A \notin \mathfrak{so}(\mathbb{A}_n)_{ST}\right\}, \tag{184}$$

where $\mathfrak{so}(\mathbb{A}_n)_{ST}$ is the subalgebra of $\mathfrak{so}(\mathbb{A}_{n\mathcal{N}})$ that acts as orthogonal transformations on $\mathbb{A}_n \subseteq \mathbb{A}_{n\mathcal{N}}$. The division algebras used in each dimension and the corresponding $\mathfrak{sym}$ algebras are summarised in Table 8.

Let us take $D = 3$ as a concrete example. The $\mathcal{N} = 8$ Lagrangian is given by

$$\begin{aligned}\mathcal{L} = \operatorname{tr}\Big(&\tfrac{1}{2}F \wedge \star F - \tfrac{1}{2}D\varphi_i \wedge \star D\varphi_i + i\bar{\lambda}_a \not{D}\lambda_a \\ &- \tfrac{1}{4}g^2[\varphi_i, \varphi_j][\varphi_i, \varphi_j] - g\bar{\lambda}^a \Gamma^i_{ab}[\varphi_i, \lambda^b]\Big),\end{aligned} \tag{185}$$

where Γ^i_{ab}, $i = 1, \ldots, 7$, $a, b = 0, \ldots, 7$, belongs to the SO(7) Clifford algebra. The key observation is that this gamma matrix can be represented by the $\mathbb{A}$ structure constants C_{abc},

$$\Gamma^i_{ab} = i(\delta_{bi}\delta_{a0} - \delta_{b0}\delta_{ai} + C_{iab}), \tag{186}$$

Table 9 Magic square of maximal compact subalgebras

$\mathbb{A}_L \backslash \mathbb{A}_R$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{O}$
$\mathbb{R}$	$\mathfrak{so}(2)$	$\mathfrak{so}(3) \times \mathfrak{so}(2)$	$\mathfrak{so}(5) \times \mathfrak{so}(3)$	$\mathfrak{so}(9)$
$\mathbb{C}$	$\mathfrak{so}(3) \times \mathfrak{so}(2)$	$[\mathfrak{so}(3) \times \mathfrak{so}(2)]^2$	$\mathfrak{so}(6) \times \mathfrak{so}(3) \times \mathfrak{so}(2)$	$\mathfrak{so}(10) \times \mathfrak{so}(2)$
$\mathbb{H}$	$\mathfrak{so}(5) \times \mathfrak{so}(3)$	$\mathfrak{so}(6) \times \mathfrak{so}(3) \times \mathfrak{so}(2)$	$\mathfrak{so}(8) \times \mathfrak{so}(4)$	$\mathfrak{so}(12) \times \mathfrak{so}(3)$
$\mathbb{O}$	$\mathfrak{so}(9)$	$\mathfrak{so}(10) \times \mathfrak{so}(2)$	$\mathfrak{so}(12) \times \mathfrak{so}(3)$	$\mathfrak{so}(16)$

which allows us to rewrite the $\mathcal{N} = 1, 2, 4, 8$ action in terms of a single expression defined over $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$:

$$\mathcal{L} = \operatorname{tr}\left(\tfrac{1}{2}F \wedge \star F - \tfrac{1}{2}D\overline{\varphi} \wedge \star D\varphi + i\bar{\lambda}\not{D}\lambda - \tfrac{1}{4}g^2\langle[\varphi, \varphi]|[\varphi, \varphi]\rangle - g[\bar{\lambda}, \varphi, \lambda]\right), \tag{187}$$

where $\varphi = \varphi^i e_i$ is an Im$\mathbb{A}$-valued scalar field, $\lambda = \lambda^a e_a$ is an $\mathbb{A}$-valued two-component spinor and $\bar{\lambda} = \bar{\lambda}^a e_a^*$.

Now consider the product of two division algebraic multiplets, where $\mathcal{N} = \dim \mathbb{A}$, $A_\mu \in \text{Re}\mathbb{A}$, $\varphi \in \text{Im}\mathbb{A}$, $\lambda \in \mathbb{A}$ and similar for the right theory. We obtain the field content of an $(\mathcal{N} + \tilde{\mathcal{N}})$-extended supergravity theory valued in both $\tilde{\mathbb{A}}$ and $\tilde{\mathbb{A}}$:

$$g_{\mu\nu} \in \mathbb{R}, \quad \Psi_\mu \in \begin{pmatrix} \mathbb{A} \\ \tilde{\mathbb{A}} \end{pmatrix}, \quad \varphi, \chi \in \begin{pmatrix} \mathbb{A} \otimes \tilde{\mathbb{A}} \\ \mathbb{A} \otimes \tilde{\mathbb{A}} \end{pmatrix}. \tag{188}$$

The $\mathbb{R}$-valued graviton and $\mathbb{A} \oplus \tilde{\mathbb{A}}$-valued gravitino carry no degrees of freedom. The $(\mathbb{A} \otimes \tilde{\mathbb{A}})^2$-valued scalar and Majorana spinor each have $2(\dim \mathbb{A} \times \dim \tilde{\mathbb{A}})$ degrees of freedom.

The $\mathcal{H}$ algebra then follows immediately in this division algebraic language. The left and right factors each come with a commuting copy of the triality algebra, $\mathfrak{tri}(\mathbb{A}) \oplus \mathfrak{tri}(\tilde{\mathbb{A}})$. However, the $\mathbb{A} \otimes \tilde{\mathbb{A}}$ doublets in (188) form irreducible representations of R-symmetry. The corresponding generators must themselves transform under $\mathfrak{tri}(\mathbb{A}) \oplus \mathfrak{tri}(\tilde{\mathbb{A}})$ consistently, implying they are elements of $\mathbb{A} \otimes \tilde{\mathbb{A}}$. This follows, formally, from the left/right supersymmetries $Q \otimes \tilde{Q}$

$$\mathfrak{h}(\mathbb{A}, \tilde{\mathbb{A}}) := \underbrace{\mathfrak{tri}(\mathbb{A})}_{\text{Left global symmetries}} \oplus \underbrace{\mathfrak{tri}(\tilde{\mathbb{A}})}_{\text{Right global symmetries}} + \underbrace{\mathbb{A} \otimes \tilde{\mathbb{A}}}_{Q \otimes \tilde{Q}}. \tag{189}$$

This follows from the observation that $Q \otimes \tilde{Q} \in \mathbb{A} \otimes \tilde{\mathbb{A}}$. Recall, these are "pseudo-supersymmetry" transformations since they do not change the mass dimension of the component fields. This Lie algebra yields the maximal compact subalgebras of the corresponding non-compact global symmetries of the magic square, as given in Table 9.

The U-dualities $\mathcal{G}$ are realised non-linearly on the scalars, which parametrise the symmetric spaces $\mathcal{G}/\mathcal{H}$. This can be understood using the identity relating $(\mathbb{A} \otimes \tilde{\mathbb{A}})^2$

to $\mathcal{G}/\mathcal{H}$,

$$(\mathbb{A}\otimes\tilde{\mathbb{A}})\mathbb{P}^2 \cong \mathcal{G}/\mathcal{H}. \tag{190}$$

The scalar fields may be regarded as points in division-algebraic projective planes. The tangent space $T_p(\mathcal{G}/\mathcal{H}) \cong \mathfrak{p} = \mathfrak{g} \ominus \mathfrak{h}$ implies the scalars carry the $\mathfrak{p}$-representation of $\mathcal{H}$. The tangent space at any point of $(\mathbb{A}\otimes\tilde{\mathbb{A}})\mathbb{P}^2$ is just $(\mathbb{A}\otimes\tilde{\mathbb{A}})^2$, the required representation space of $\mathcal{H}$. Since $\mathcal{G}/\mathcal{H}$ is a symmetric space, the U-duality Lie algebra is given by adjoining the scalar representation space $(\mathbb{A}\otimes\tilde{\mathbb{A}})^2$ to (189),

$$\mathfrak{ms}(\mathbb{A},\tilde{\mathbb{A}}) := \underbrace{\mathfrak{tri}(\mathbb{A})\oplus\mathfrak{tri}(\tilde{\mathbb{A}}) + (\mathbb{A}\otimes\tilde{\mathbb{A}})}_{\mathfrak{h}(\mathbb{A},\tilde{\mathbb{A}})} + \underbrace{(\mathbb{A}\otimes\tilde{\mathbb{A}})^2}_{\text{“scalars”}}. \tag{191}$$

This has a $\mathbb{Z}_2\times\mathbb{Z}_2$ graded Lie algebra structure uniquely determined by the left/right super Yang–Mills factors and yields precisely the magic square [138].

Let us describe how this formula works. For an element $A \in \mathfrak{so}(\mathbb{A})$ define $\sigma A \equiv \tau A \tau^{-1} \in \mathfrak{so}(\mathbb{A})$. Then

$$\theta : \mathfrak{tri}(\mathbb{A}) \to \mathfrak{tri}(\mathbb{A}) : (A, B, C) \mapsto (\sigma B, C, \sigma A), \tag{192}$$

is an order three Lie algebra automorphism, which for $\mathbb{A} = \mathbb{O}$ interchanges the three inequivalent 8-dimensional representations of $\mathfrak{so}(\mathbb{O})$.

Given two normed division algebras $\mathbb{A}$ and $\tilde{\mathbb{A}}$ we can define on

$$\mathfrak{ms}(\mathbb{A},\tilde{\mathbb{A}}) = [\mathfrak{tri}(\mathbb{A})\oplus\mathfrak{tri}(\tilde{\mathbb{A}})]_{00} + (\mathbb{A}\otimes\tilde{\mathbb{A}})_{01} + (\mathbb{A}\otimes\tilde{\mathbb{A}})_{10} + (\mathbb{A}\otimes\tilde{\mathbb{A}})_{11} \tag{193}$$

a $\mathbb{Z}_2\times\mathbb{Z}_2$ graded Lie algebra structure. First, $\mathfrak{tri}(\mathbb{A})$ and $\mathfrak{tri}(\tilde{\mathbb{A}})$ are Lie subalgebras. For elements $T = (A, \sigma B, \sigma C)$ in $\mathfrak{tri}(\mathbb{A})$ and $(a\otimes b, 0, 0)$, $(0, a\otimes b, 0)$, and $(0, 0, a\otimes b)$ in $3(\mathbb{A}\otimes\tilde{\mathbb{A}})$, the commutators are given by the natural action of $\mathfrak{tri}(\mathbb{A})$,

$$\begin{aligned}
[T, (a\otimes b, 0, 0)] &= (A(a)\otimes b, 0, 0),\\
[T, (0, a\otimes b, 0)] &= (0, B(a)\otimes b, 0),\\
[T, (0, 0, a\otimes b)] &= (0, 0, C(a)\otimes b).
\end{aligned} \tag{194}$$

Similarly for $\tilde{T} = (\tilde{A}, \sigma\tilde{B}, \sigma\tilde{C})$ in $\mathfrak{tri}(\tilde{\mathbb{A}})$,

$$\begin{aligned}
[\tilde{T}, (a\otimes b, 0, 0)] &= (a\otimes\tilde{A}(b), 0, 0),\\
[\tilde{T}, (0, a\otimes b, 0)] &= (0, a\otimes\tilde{B}(b), 0),\\
[\tilde{T}, (0, 0, a\otimes b)] &= (0, 0, a\otimes\tilde{C}(b)).
\end{aligned} \tag{195}$$

For two elements belonging to the same summand $(\mathbb{A}\otimes\tilde{\mathbb{A}})_{ij}$ in (193) the commutators are given by

$$\begin{aligned}
[(a\otimes b, 0, 0), (a'\otimes b', 0, 0)] &= \ \ \langle a, a'\rangle\ \tilde{T}_{b,b'} + \langle b, b'\rangle T_{a,a'},\\
[(0, a\otimes b, 0), (0, a'\otimes b', 0)] &= -\langle a, a'\rangle\ \theta\tilde{T}_{b,b'} - \langle b, b'\rangle \theta T_{a,a'},
\end{aligned}$$

$$[(0,0,a\otimes b),(0,0,a'\otimes b')] = -\langle a,a'\rangle\theta^2\tilde{T}_{b,b'} - \langle b,b'\rangle\theta^2 T_{a,a'}, \tag{196}$$

where

$$T_{a,a'} := (S_{a,a'}, R_{a'}R_{\overline{a}} - R_a R_{\overline{a}'}, L_{a'}L_{\overline{a}} - L_a L_{\overline{a}'}), \tag{197}$$

and

$$S_{a,a'}(b) = \langle a,b\rangle a' - \langle a',b\rangle a, \quad L_a(b) = ab, \quad R_a(b) = ba. \tag{198}$$

Finally, we have

$$\begin{aligned}
[(a\otimes b,0,0),(0,a'\otimes b',0)] &= (0,0,\overline{aa'}\otimes\overline{bb'}),\\
[(0,0,a\otimes b),(a'\otimes b',0,0)] &= (0,\overline{aa'}\otimes\overline{bb'},0),\\
[(0,a\otimes b,0),(0,0,a'\otimes b')] &= -(\overline{aa'}\otimes\overline{bb'},0,0).
\end{aligned} \tag{199}$$

With these commutators the magic square formula (193) describes the Lie algebras of Table 6. The formula (191) is based on the triality construction described in [196]. Although isomorphic as vector spaces, they have different Lie algebra structures, as reflected in the distinct real forms appearing in each case. We see that we truncate to the maximal compact subalgebra (189) by discarding any two of the three summands $(\mathbb{A}\otimes\tilde{\mathbb{A}})_{ij}$.

For $D = n+2$, we begin with a pair of Yang–Mills theories with $\mathcal{N}$ and $\tilde{\mathcal{N}}$ supersymmetries written over the division algebras $\mathbb{A}_{n\mathcal{N}}$ and $\mathbb{A}_{n\tilde{\mathcal{N}}}$, respectively, as described in [190]. In terms of spacetime little group representations we may then write all the bosons of the left (right) theory as a single element $b\in\mathbb{A}_{n\mathcal{N}}$ ($\tilde{b}\in\mathbb{A}_{n\tilde{\mathcal{N}}}$), and similarly for the fermions $f\in\mathbb{A}_{n\mathcal{N}}$ ($\tilde{f}\in\mathbb{A}_{n\tilde{\mathcal{N}}}$). After tensoring we arrange the resulting supergravity fields into a bosonic doublet and a fermionic doublet,

$$B = \begin{pmatrix} b\otimes\tilde{b}\\ f\otimes\tilde{f}\end{pmatrix}, \quad F = \begin{pmatrix} b\otimes\tilde{f}\\ f\otimes\tilde{b}\end{pmatrix}, \tag{200}$$

just as we did in $D = 3$. The algebra (189) acts naturally on these doublets. However, a diagonal $\mathfrak{so}(\mathbb{A}_n)_{ST}$ subalgebra of this corresponds to spacetime transformations, so we must restrict $\mathfrak{h}(\mathbb{A}_{n\mathcal{N}},\mathbb{A}_{n\tilde{\mathcal{N}}})$ to the subalgebra that commutes with $\mathfrak{so}(\mathbb{A}_n)_{ST}$. Heuristically, we identify a diagonal spacetime subalgebra $\mathbb{A}_n$ in $\mathbb{A}_{n\mathcal{N}}\otimes\mathbb{A}_{n\tilde{\mathcal{N}}}$ and require that it is preserved by the global isometries, which picks out a subset in $\mathfrak{Isom}((\mathbb{A}_{n\mathcal{N}}\otimes\mathbb{A}_{n\tilde{\mathcal{N}}})\mathbb{P}^2)$. Imposing this condition selects the U-duality algebra of the $D = n+2$, $(\mathcal{N}+\tilde{\mathcal{N}})$-extended supergravity theory obtained by tensoring $\mathcal{N}$ and $\tilde{\mathcal{N}}$ super Yang–Mills theories. The Lie algebras are given by the *magic pyramid formula*:

$$\mathfrak{mp}(\mathbb{A}_n,\mathbb{A}_{n\mathcal{N}},\mathbb{A}_{n\tilde{\mathcal{N}}}) := \left\{ u\in\mathfrak{m}(\mathbb{A}_{n\mathcal{N}},\mathbb{A}_{n\tilde{\mathcal{N}}})\ominus\mathfrak{so}(\mathbb{A}_n)_{ST} \,\middle|\, [u,\mathfrak{so}(\mathbb{A}_n)_{ST}] = 0\right\}. \tag{201}$$

The terminology is made clear by the pyramid of corresponding U-dualities groups presented in Fig. 1. The base of the pyramid in $D = 3$ is the 4×4 Freudenthal magic square, while the higher levels are comprised of a 3×3 square in $D = 4$, a 2×2 square in $D = 6$ and Type II supergravity at the apex in $D = 10$. Note, in [313] the oxidation of $\mathcal{N}$-extended $D = 3$ dimensional supergravity theories was shown to generate a partially symmetric "trapezoid" of non-compact global symmetries for $D = 3, 4, \ldots 11$ and $0, 2^0, 2^1, \ldots 2^7$ supercharges. A subset of algebras in the trapezoid with $D = 3, 4, 5$ and $2^5, 2^6, 2^7$ supercharges matches the $D = 3, 4, 5$ and $\mathbb{A} = \mathbb{C}, \mathbb{H}, \mathbb{O}$ exterior wall of the pyramid of Fig. 1.

To illustrate the principles of the magic pyramid let us consider the simplest example in $D = 4$, the product of two $\mathcal{N} = 1$ Yang–Mills multiplets (A_μ, λ), which must yield $\mathcal{N} = 2$ supergravity coupled to one hypermultiplet. This follows from state counting and supersymmetry alone, but the actual coupling is not fixed. By determining the symmetry this ambiguity is resolved (assuming, as before, a homogenous scalar manifold). The left/right Yang–Mills on-shell multiplets are represented by the complex numbers (helicity states):

$$A,\ \lambda \in \mathbb{C}, \quad \tilde{A},\ \tilde{\lambda} \in \tilde{\mathbb{C}}. \tag{202}$$

Collecting the bosonic/fermionic states, the product gives us the $(\mathbb{C} \otimes \mathbb{C})^2$ valued objects:

$$B = \begin{pmatrix} A \otimes \tilde{A} \\ \lambda \otimes \tilde{\lambda} \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} A \otimes \tilde{\lambda} \\ \lambda \otimes \tilde{A} \end{pmatrix}. \tag{203}$$

Let us consider the $D = 3$ maximal compact algebra

$$\mathfrak{h}(\mathbb{C}, \tilde{\mathbb{C}}) := \mathfrak{tri}(\mathbb{C}) \oplus \mathfrak{tri}(\tilde{\mathbb{C}}) + \mathbb{C} \otimes \tilde{\mathbb{C}}, \tag{204}$$

in this representation. To describe the generators acting on $(\mathbb{C} \otimes \mathbb{C})^2$ it is convenient to define the quantities

$$1_\pm := \frac{1}{2}(1 \otimes 1 \mp i \otimes i), \quad i_\pm := \frac{1}{2}(i \otimes 1 \pm 1 \otimes i), \tag{205}$$

which form two orthogonal copies[37] of $\mathbb{C}$:

$$1_\pm^2 = 1_\pm, \ \ 1_\pm i_\pm = i_\pm, \ \ , i_\pm^2 = -1_\pm \ \ 1_\pm 1_\mp = 0, \ \ 1_\pm i_\mp = 0, \ \ i_\pm i_\mp = 0 \tag{206}$$

A basis for $\mathfrak{tri}(\mathbb{C}) \cong \mathfrak{so}(2) \oplus \mathfrak{so}(2)$ is given by the $\mathbb{C} \otimes \tilde{\mathbb{C}}$-valued 2×2 matrices

$$i_+ \mathbb{1}, \quad i_+ \sigma^1, \tag{207}$$

while those of $\mathfrak{tri}(\tilde{\mathbb{C}})$ are given by $i_- \mathbb{1},\ i_- \sigma^1$. The generators of the $\mathbb{C} \otimes \tilde{\mathbb{C}}$ term are similarly given by

$$1_+ \varepsilon, \quad i_+ \sigma^3, \quad 1_- \varepsilon, \quad i_- \sigma^3. \tag{208}$$

It is straightforwd to verify that these matrices generate $\mathfrak{su}(2) \times \mathfrak{su}(2) \times \mathfrak{u}(1) \times \mathfrak{u}(1) \cong \mathfrak{so}(4) \times \mathfrak{so}(2) \times \mathfrak{so}(2)$, as stated in the compact sub-magic square in Table 9.

---

[37] The objects $1_\pm$ act as projection operators dividing $\mathbb{C} \otimes \mathbb{C}$ into two 2-dimensional subspaces, on which $i_\pm$ act as complex structures, so that $\mathbb{C} \otimes \mathbb{C} \cong \mathbb{C} \oplus \mathbb{C}$.

Thus far this is just the $D = 3$ analysis. But recall, in $D = 4$ a diagonal $\mathbb{C} \in \mathbb{C} \otimes \tilde{\mathbb{C}}$ is identified with spacetime. The left Yang–Mills multiplets transform under its spacetime little algebra $\mathfrak{u}(1)$ acting on $\mathbb{C}$ as

$$\delta_\theta A = i\theta A, \qquad \delta_\theta \lambda = \frac{1}{2} i\theta\lambda, \tag{209}$$

(together with the complex conjugates) and similar for the right multiplet, with $\tilde{\mathfrak{u}}(1)$ acting on $\tilde{\mathbb{C}}$ with parameter $\tilde{\theta}$. Focussing on the fermion doublet and identifying the left/right spacetimes, $\theta = \tilde{\theta}$, we find

$$\delta_\theta F = \theta \left(\frac{3}{2} i_+ \mathbb{1} + \frac{1}{2} i_- \sigma^3 \right) F. \tag{210}$$

Let us unravel what is happening here. Focusing on the fermions we see that the positive helicity spin-$\frac{3}{2}$ and spin-$\frac{1}{2}$ states belong to i_+ and i_- sectors, just as one would anticipate. This is only consistent because $\mathbb{C} \otimes \mathbb{C}$ is *not* a division algebra; it contains zero divisors and we have $i_+ i_- = 0$. Their role of the failure of the division property here is to ensure that each component in the multiplet transforms with the correct helicity and only the correct helicity. Having identified the spacetime little group generator, the remaining internal symmetries are determined. All the matrices commute with $i_+\mathbb{1}$, but $i_-\sigma^1$ and $1_-\epsilon$ do not commute with $i_-\sigma^3$, so we are forced to discard these generators, leaving the subalgebra

$$\mathfrak{u}(1) \oplus \mathfrak{u}(1) \oplus \mathfrak{su}(2). \tag{211}$$

This is the maximal compact subalgebra of the corresponding $D = 4, \mathcal{N} = 2$ (or $\mathbb{C}, \mathbb{C}, \mathbb{C}$) entry in the pyramid, as given in Table 9. Acting on the gravitino with these generators we find it transforms as a doublet but, again because of the i_+ annihilating i_-, the spin-$\frac{1}{2}$ fields are singlets, as required in the supergravity theory. The Yang–Mills R-symmetries have been absorbed into the U-duality group. A similar analysis for the bosonic fields in the theory shows that we do indeed obtain a graviton, a vector and two scalars, which transform as a singlet, a singlet and doublet under the $\mathfrak{su}(2)$, as required.

Applying the same prinicples to the non-compact global symmetries we arrive at the complete pyramid given in Fig. 1. Let us just illustrate the point by returning to our discussion of $\mathcal{N} = 8$ supergravity in $D = 4$, which we expect to yield $E_{7(7)}$. Just as for the case treated above we need to identify the $\mathbb{C} \subset \mathbb{O} \otimes \tilde{\mathbb{O}}$ corresponding to the $D = 4$ spacetime algebra. To do so we start from the $D = 3$ $(\mathbb{O}, \mathbb{O})$ entry of the magic square and decompose with respect to left/right Yang–Mills symmetries:

$$\begin{aligned}
&\mathfrak{e}_{8(8)} \supset \mathfrak{so}(8) \oplus \mathfrak{so}(8) \supset \mathfrak{su}(4) \oplus \mathfrak{u}(1) \oplus \mathfrak{su}(4) \oplus \mathfrak{u}(1) \\
&\mathbf{248} \to \big[(\mathbf{15},\mathbf{1})_{00} + (\mathbf{1},\mathbf{1})_{00} \quad + (\mathbf{1},\mathbf{15})_{00} + (\mathbf{1},\mathbf{1})_{00} \\
&\qquad + (\mathbf{6},\mathbf{1})_{20} + (\mathbf{1},\mathbf{6})_{02} \quad + (\mathbf{1},\mathbf{6})_{0-2} + (\mathbf{6},\mathbf{1})_{-20} \\
&\qquad + (\mathbf{4},\mathbf{4})_{11} + (\mathbf{4},\bar{\mathbf{4}})_{1-1} + (\bar{\mathbf{4}},\mathbf{4})_{-11} + (\bar{\mathbf{4}},\bar{\mathbf{4}})_{-1-1} \big]
\end{aligned}$$

$$
\begin{aligned}
&+ \big[(\overline{\mathbf{4}},\overline{\mathbf{4}})_{11} + (\overline{\mathbf{4}},\mathbf{4})_{1-1} + (\mathbf{4},\overline{\mathbf{4}})_{-11} + (\mathbf{4},\mathbf{4})_{-1-1} \\
&+ (\mathbf{1},\mathbf{1})_{22} + (\mathbf{1},\mathbf{1})_{2-2} + (\mathbf{1},\mathbf{1})_{-22} + (\mathbf{1},\mathbf{1})_{-2-2} \\
&+ (\mathbf{1},\mathbf{6})_{20} + (\mathbf{6},\mathbf{1})_{02} \quad + (\mathbf{6},\mathbf{1})_{0-2} + (\mathbf{1},\mathbf{6})_{-20} + (\mathbf{6},\mathbf{6})_{00}\big],
\end{aligned} \tag{212}
$$

where we have separated the compact ($\mathfrak{tri}(\mathbb{O}) \oplus \mathfrak{tri}(\mathbb{O}) + \mathbb{O} \otimes \tilde{\mathbb{O}}$) and non-compact ($\mathbb{O}\otimes\tilde{\mathbb{O}}+\mathbb{O}\otimes\tilde{\mathbb{O}}$) generators with the square parentheses. To distinguish the spacetime little group $\mathfrak{u}(1)_{\mathrm{st}}$ from the internal $\mathfrak{u}(1)$ we must take the sum and difference of the $\mathfrak{u}(1) \oplus \mathfrak{u}(1)$ charges giving,

$$
\begin{aligned}
&\mathfrak{e}_{8(8)} \supset \mathfrak{so}(8) \oplus \mathfrak{so}(8) \supset \mathfrak{u}(1)_{\mathrm{st}} \oplus \mathfrak{su}(4) \oplus \mathfrak{su}(4) \oplus \mathfrak{u}(1) \\
&\mathbf{248} \to \big[(\mathbf{15},\mathbf{1})_{00} + (\mathbf{1},\mathbf{1})_{00} + (\mathbf{1},\mathbf{15})_{00} + (\mathbf{1},\mathbf{1})_{00} \\
&\quad + (\mathbf{6},\mathbf{1})_{22} + (\mathbf{1},\mathbf{6})_{2-2} + (\mathbf{1},\mathbf{6})_{-22} + (\mathbf{6},\mathbf{1})_{-2-2} \\
&\quad + (\mathbf{4},\mathbf{4})_{20} + (\mathbf{4},\overline{\mathbf{4}})_{02} + (\overline{\mathbf{4}},\mathbf{4})_{0-2} + (\overline{\mathbf{4}},\overline{\mathbf{4}})_{-20}\big] \\
&\quad \times \big[+ (\overline{\mathbf{4}},\overline{\mathbf{4}})_{20} + (\overline{\mathbf{4}},\mathbf{4})_{02} + (\mathbf{4},\overline{\mathbf{4}})_{0-2} + (\mathbf{4},\mathbf{4})_{-20} \\
&\quad + (\mathbf{1},\mathbf{1})_{40} + (\mathbf{1},\mathbf{1})_{04} + (\mathbf{1},\mathbf{1})_{0-4} + (\mathbf{1},\mathbf{1})_{-40} \\
&\quad + (\mathbf{1},\mathbf{6})_{22} + (\mathbf{6},\mathbf{1})_{2-2} + (\mathbf{6},\mathbf{1})_{-22} + (\mathbf{1},\mathbf{6})_{-2-2} + (\mathbf{6},\mathbf{6})_{00}\big],
\end{aligned} \tag{213}
$$

where the first charge corresponds to $\mathfrak{u}(1)_{\mathrm{st}}$. The $D = 4$ global symmetry generators are those left once we have discarded all generators carrying a non-trivial $\mathfrak{u}(1)_{\mathrm{st}}$ charge (as well as the $\mathfrak{u}(1)_{\mathrm{st}}$ itself, of course), which yields

$$
\begin{aligned}
\mathbf{133} \to &\big[(\mathbf{1},\mathbf{1})_0 + (\mathbf{15},\mathbf{1})_0 + (\mathbf{1},\mathbf{15})_0 + (\mathbf{4},\overline{\mathbf{4}})_2 + (\overline{\mathbf{4}},\mathbf{4})_{-2}\big] \\
&+ \big[(\mathbf{4},\overline{\mathbf{4}})_{-2} + (\overline{\mathbf{4}},\mathbf{4})_2 + (\mathbf{1},\mathbf{1})_4 + (\mathbf{1},\mathbf{1})_{-4} + (\mathbf{6},\mathbf{6})_0\big].
\end{aligned} \tag{214}
$$

We recognise this as precisely the decomposition of $\mathfrak{e}_{7(7)}$ under $\mathfrak{e}_{7(7)} \supset \mathfrak{su}(8) \supset \mathfrak{su}(4) \oplus \mathfrak{su}(4) \oplus \mathfrak{u}(1)$: where the compact pieces, contained in the first bracket, form the maximal compact subalgebra $\mathfrak{su}(8)$,

$$
\begin{aligned}
&\mathrm{SU}(8) \supset \mathrm{SU}(4) \times \mathrm{SU}(4) \times \mathrm{Un}(1) \\
&\quad \mathbf{63} \to (\mathbf{1},\mathbf{1})_0 + (\mathbf{15},\mathbf{1})_0 + (\mathbf{1},\mathbf{15})_0 + (\mathbf{4},\overline{\mathbf{4}})_2 + (\overline{\mathbf{4}},\mathbf{4})_{-2}.
\end{aligned} \tag{215}
$$

To extract the field content we simply decompose the **128** (B) and **128**′ (F) of $\mathfrak{so}(16)$ with respect to $\mathfrak{u}(1)_{\mathrm{st}} \oplus \mathfrak{su}(8)$,

$$
\begin{aligned}
\mathbf{128} &\to \mathbf{1}_2 + \mathbf{1}_{-2} + \mathbf{28}_1 + \overline{\mathbf{28}}_{-1} + \mathbf{70}_0 \\
\mathbf{128}' &\to \mathbf{8}_3 + \overline{\mathbf{8}}_{-3} + \mathbf{56}_1 + \overline{\mathbf{56}}_{-1},
\end{aligned} \tag{216}
$$

which yields the helicity states and global representations of the expected $\mathcal{N} = 8$ supermultiplet.

Let us conclude with some comments on the product of theories other than super Yang–Mills. Particularly interesting examples are provided by the superconformal multiplets in $D = 3, 4, 6$. In a manner directly analogous to the magic pyramid the tensor product of left and right superconformal theories yields the "conformal pyramid", described in [138]. It has the remarkable property that its faces are also given by the Freudenthal magic square. In particular, ascending up the maximal spine one encounters the famous exceptional sequence $E_{8(8)}, E_{7(7)}, E_{6(6)}$, but where $E_{6(6)}$ belongs to the $D = 6$, (4, 0) theory proposed by Hull as the superconformal limit of M-theory compactified on a 6-torus [276,366,367]. This pattern suggests the existence of some highly exotic $D = 10$ theory with $F_{4(4)}$ U-duality group. The existence of such a theory would be more than a little surprising and there is a (slightly) more conventional interpretation of the conformal pyramid, including its $F_{4(4)}$ tip, but for theories in $D = 3, 4, 5, 6$, as described [138].

The product of conformal theories in the context of amplitudes has been considered previously in, for example, [66,70,71,368,369]. In particular, the maximally supersymmetric $D = 3, \mathcal{N} = 8$ Bagger-Lambert-Gustavsson (BLG) Chern-Simons-matter theory [210–212] has been shown to enjoy a colour-kinematic duality reflecting its three-algebra structure [71]. The "square" of BLG amplitudes yields those of $\mathcal{N} = 16$ supergravity. Since $\mathcal{N} = 16$ supergravity is the unique theory with 32 supercharges in three dimensions it is also the "square" of the $\mathcal{N} = 8$ Yang–Mills theory. The square of the amplitudes in both cases agree, despite their distinct structures [70].

In $D = 6$ one might expect relations between the "square" of the $\mathcal{N} = (2, 0)$ tensor multiplet and the $\mathcal{N} = (4, 0)$ theory proposed by Hull [276,366,367], as put forward in [66]. Of course, amplitudes are generically not well-defined in these cases, but one can make some precise statements in terms of the tree-level S-matrix in particular regimes, as discussed in [368,369]. For example, in the absence of additional degrees of freedom all three-point tree-level amplitudes of the (2, 0) tensor multiplet vanish [368]. The $D = 5, \mathcal{N} = 4$ super Yang–Mills theory squares to give the amplitudes of $D = 5, \mathcal{N} = 8$ supergravity. However, being non-renormalisable it ought to be regarded as a superconformal $D = 6, \mathcal{N} = (2, 0)$ theory compactified on a circle of radius $R = g^2_{YM}/4\pi^2$. At linearised level Hull's (4, 0) theory follows from the square the (2, 0) theory [142] and gives $\mathcal{N} = 8, D = 5$ supergravity when compactified on a circle [366]. In fact, it would seem that one can go beyond the free theory and construct candidate tree-level amplitudes of the (4, 0) theory from the double-copy of the tree-level amplitudes of the (2, 0) theory using $D = 6$ spinor-helicities and a polarised version of the Cachazo-He-Yuan scattering equation formalism [269]. They necessarily start at four-points, since the three-point (2, 0) amplitudes are trivial. Of course, the (4, 0) amplitudes need to be tested before one can claim they correspond to Hull's conjecture. For instance, their double-soft scalar limits should reveal the $E_{6(6)}$ symmetry. They do already pass the first test; on dimensional reduction on a circle they yield the amplitudes of $D = 5, \mathcal{N} = 8$ supergravity. From this perspective the $(2, 0) \times (2, 0) = (4, 0)$ identity constitutes an, as yet ill-defined, M-theory up-lift of the maximally supersymmetric $D = 5$ squaring relation.

4 Closing remarks

We started our journey by posing a number of questions regarding the nature of the "gravity = gauge × gauge" paradigm. In particular, we asked: (1) Why does the correspondence work? (2) Is it strictly a property of amplitudes or can it be generalised to other/all aspects of gauge and gravity theories? (3) What classes of gravitational theories admit a gauge theory squared origin?

In the course of the subsequent discussion we have witnessed remarkable progress on all fronts. Our understanding of BCJ duality and the double-copy has developed dramatically and along with it our handle on the divergences of perturbative quantum gravity. The programme has clearly shown itself to be an effective point of view beyond amplitudes from a number of perspectives, from novel approaches to the construction of solutions to the identification of new gauge and gravity theories. It has also become increasingly clear that the BCJ double-copy, and "gravity $=$ gauge $\times$ gauge" more generally, can be applied to a vast and diverse set of theories, that continues to grow.

Yet, we have no complete answers and the central questions remain: Does BCJ duality hold to all orders, at least for some theories? What are the ultimate implications for quantum gravity? Is there a geometrical or world-sheet underpinning? Can we characterise all theories admitting a gauge squared origin? Is it a/the right way to think of gravity?

Acknowledgements I gratefully acknowledge T. Adamo, Z. Bern, B. Cerchiai, M. Chiodaroli, S. Deser, S. Ferrara, Y. Geyer, A. Hodges, C. M. Hull, H. Johansson, B. Julia, A. Luna, L. Mason, G. Mogull, R. Monteiro, A. Ochirov, D. O'Connell, R. Roiban, O. Schlotterer, W. Siegel, K. Stelle and C. White and for many helpful conversations and correspondences. I would especially like to thank my collaborators A. Anastansiou, M. J. Duff, M. J. Hughes, A. Marrani, S. Nagy and M. Zoccali. This work was supported by a Schrödinger Fellowship.

References

1. S.W. Hawking, Particle creation by black holes. Commun. Math. Phys. **43**, 199–220 (1975)
2. S.W. Hawking, Black holes and thermodynamics. Phys. Rev. D **13**, 191–197 (1976)
3. G. 't Hooft, M.J.G. Veltman, One loop divergencies in the theory of gravitation. Ann. Inst. His. Poincare Phys. Theor. A **20**, 69–94 (1974)
4. S. Deser, P. van Nieuwenhuizen, Nonrenormalizability of the quantized Einstein–Maxwell system. Phys. Rev. Lett. **32**, 245–247 (1974)
5. S. Deser, P. van Nieuwenhuizen, Nonrenormalizability of the quantized Dirac–Einstein system. Phys. Rev. D **10**, 411 (1974)
6. M.H. Goroff, A. Sagnotti, Quantum gravity at two loops. Phys. Lett. B **160**, 81–86 (1985)
7. M.H. Goroff, A. Sagnotti, The ultraviolet behavior of Einstein gravity. Nucl. Phys. B **266**, 709–736 (1986)
8. T. Kaluza, Zum Unitätsproblem der Physik. Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1921**, 966–972 (1921). arXiv:1803.08616 [physics.hist-ph]. [Int. J. Mod. Phys. D **27**(14), 1870001 (2018)]
9. O. Klein, Quantum theory and five-dimensional theory of relativity. (In German and English). Z. Phys. **37**, 895–906 (1926). [76(1926)]
10. O. Klein, The atomicity of electricity as a quantum theory law. Nature **118**, 516 (1926)
11. M.J. Duff, B.E.W. Nilsson, C.N. Pope, Kaluza–Klein supergravity. Phys. Rept. **130**, 1–142 (1986)
12. R. Utiyama, Invariant theoretical interpretation of interaction. Phys. Rev. **101**, 1597–1607 (1956)
13. T.W.B. Kibble, Lorentz invariance and the gravitational field. J. Math. Phys. **2**, 212–221 (1961)
14. S.W. MacDowell, F. Mansouri, Unified geometric theory of gravity and supergravity. Phys. Rev. Lett. **38**, 739–742 (1977)
15. A.H. Chamseddine, P.C. West, Supergravity as a gauge theory of supersymmetry. Nucl Phys. B **129**, 39–44 (1977)

16. K.S. Stelle, P.C. West, de Sitter gauge invariance and the geometry of the Einstein–Cartan theory. J. Phys. A Math. Gen. **12**, L205–L210 (1979)
17. S. W. MacDowell, F. Mansouri, Unified geometric theory of gravity and supergravity. Phys. Rev. Lett. **38**, 739 (1977). [Erratum: Phys. Rev. Lett. **38**, 1376 (1977)]
18. M. Kaku, P. K. Townsend, P. van Nieuwenhuizen, Properties of conformal supergravity. Phys. Rev. D **17**, 3179 (1978). [853(1978)]
19. K.S. Stelle, P.C. West, Spontaneously broken De Sitter symmetry and the gravitational holonomy group. Phys. Rev. D **21**, 1466 (1980)
20. G. Hooft, Dimensional reduction in quantum gravity. Salamfestschr. Collect. Talks **4**(A), 1–13 (1994). arXiv:gr-qc/9310026
21. L. Susskind, The World as a hologram. J. Math. Phys. **36**, 6377–6396 (1995). arXiv:hep-th/9409089 [hep-th]
22. J.M. Maldacena, The large N limit of superconformal field theories and supergravity. Adv. Theor. Math. Phys. **2**, 231–252 (1998). arXiv:hep-th/9711200
23. S. Gubser, I.R. Klebanov, A.M. Polyakov, Gauge theory correlators from noncritical string theory. Phys. Lett. B **428**, 105–114 (1998). arXiv:hep-th/9802109 [hep-th]
24. E. Witten, Anti-de Sitter space and holography. Adv. Theor. Math. Phys. **2**, 253–291 (1998). arXiv:hep-th/9802150 [hep-th]
25. D.E. Berenstein, J.M. Maldacena, H.S. Nastase, Strings in flat space and pp waves from N = 4 super Yang–Mills. JHEP **04**, 013 (2002). arXiv:hep-th/0202021 [hep-th]
26. G. 't Hooft, A planar diagram theory for strong interactions. Nucl. Phys. B **72**, 461 (1974). [337(1973)]
27. G.T. Horowitz, J. Polchinski, Gauge/gravity duality, in *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter*, ed. by D. Oriti (Cambridge University Press, Cambridge, 2006), pp. 169–186. arXiv:gr-qc/0602037 [gr-qc]
28. V.E. Hubeny, The AdS/CFT correspondence. Class. Quant. Gravit. **32**(12), 124010 (2015). arXiv:1501.00007 [gr-qc]
29. P. Jordan, Zur neutrinotheorie des lichtes. Zeitschrift für Physik **93**(7–8), 464–472 (1935)
30. R.P. Feynman, F.B. Morinigo, W.G. Wagner, *Feynman Lectures on Gravitation* (CRC Press, New York, 2018)
31. G. Papini, Photon bound states and gravitons. Il Nuovo Cimento (1955–1965) **39**, 716–726 (1965). https://doi.org/10.1007/BF02735838
32. H. Terazawa, Y. Chikashige, K. Akama, T. Matsuki, Simple relation between the fine structure and gravitational constants. Phys. Rev. D **15**, 1181 (1977)
33. S. Weinberg, E. Witten, Limits on massless particles. Phys. Lett. B **96**, 59–62 (1980)
34. H. Kawai, D. Lewellen, S. Tye, A relation between tree amplitudes of closed and open strings. Nucl. Phys. B **269**, 1 (1986)
35. M.B. Green, J.H. Schwarz, E. Witten, Superstring theory vol. 1: Introduction. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, p. 469 (1987)
36. L. Brink, J.H. Schwarz, J. Scherk, Supersymmetric Yang–Mills theories. Nucl. Phys. B **121**, 77–92 (1977)
37. M.A. Virasoro, Alternative constructions of crossing-symmetric amplitudes with regge behavior. Phys. Rev. **177**, 2309–2311 (1969)
38. J.A. Shapiro, Narrow-resonance model with regge behavior for pi pi scattering. Phys. Rev. **179**, 1345–1353 (1969)
39. G. Veneziano, Construction of a crossing—symmetric, Regge behaved amplitude for linearly rising trajectories. Nuovo Cim. **A57**, 190–197 (1968)
40. Z. Bern, D.C. Dunbar, T. Shimada, String based methods in perturbative gravity. Phys. Lett. B **312**, 277–284 (1993). arXiv:hep-th/9307001 [hep-th]
41. Z. Bern, L.J. Dixon, D.C. Dunbar, D.A. Kosower, One loop n point gauge theory amplitudes, unitarity and collinear limits. Nucl. Phys. B **425**, 217–260 (1994). arXiv:hep-ph/9403226 [hep-ph]
42. Z. Bern, L.J. Dixon, D.C. Dunbar, D.A. Kosower, Fusing gauge theory tree amplitudes into loop amplitudes. Nucl. Phys. B **435**, 59–101 (1995). arXiv:hep-ph/9409265 [hep-ph]
43. Z. Bern, A.G. Morgan, Massive loop amplitudes from unitarity. Nucl. Phys. B **467**, 479–509 (1996). arXiv:hep-ph/9511336 [hep-ph]
44. L.J. Dixon, Calculating scattering amplitudes efficiently. In: QCD and beyond, proceedings, theoretical advanced study institute in elementary particle physics, TASI-95, Boulder, pp. 539–

584 (1996). arXiv:hep-ph/9601359 [hep-ph]. http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-PUB-7106

45. Z. Bern, L.J. Dixon, D.A. Kosower, Progress in one loop QCD computations. Ann. Rev. Nucl. Part. Sci. **46**, 109–148 (1996). arXiv:hep-ph/9602280 [hep-ph]
46. R. Britto, Loop amplitudes in gauge theories: modern analytic approaches. J. Phys. A **44**, 454006 (2011). arXiv:1012.4493 [hep-th]
47. Z. Bern, Y.-T. Huang, Basics of generalized unitarity. J. Phys. A **44**, 454003 (2011). arXiv:1103.1869 [hep-th]
48. Z. Bern, L.J. Dixon, D.C. Dunbar, M. Perelstein, J.S. Rozowsky, On the relationship between Yang–Mills theory and gravity and its implication for ultraviolet divergences. Nucl. Phys. B **530**, 401–456 (1998). arXiv:hep-th/9802162 [hep-th]
49. H. Elvang, Y.-T. Huang, Scattering amplitudes in gauge theory and gravity. Cambridge University Press, Cambridge (2015). http://www.cambridge.org/mw/academic/subjects/physics/theoretical-physics-and-mathematical-physics/scattering-amplitudes-gauge-theory-and-gravity?format=AR
50. Z. Bern, J. Carrasco, H. Johansson, New relations for gauge-theory amplitudes. Phys. Rev. D **78**, 085011 (2008). arXiv:0805.3993 [hep-ph]
51. Z. Bern, J.J.M. Carrasco, H. Johansson, Perturbative quantum gravity as a double copy of gauge theory. Phys. Rev. Lett. **105**, 061602 (2010). arXiv:1004.0476 [hep-th]
52. M. Kiermaier, Gravity as the square of gauge theory. Amplitudes 2010, Queen Mary, University of London (2010). https://strings.ph.qmul.ac.uk/~theory/Amplitudes2010/Talks/MK2010.pdf
53. N.E.J. Bjerrum-Bohr, P.H. Damgaard, T. Sondergaard, P. Vanhove, The momentum kernel of gauge and gravity theories. JHEP **01**, 001 (2011). arXiv:1010.3933 [hep-th]
54. C.R. Mafra, O. Schlotterer, S. Stieberger, Explicit BCJ numerators from pure spinors. JHEP **07**, 092 (2011). arXiv:1104.5224 [hep-th]
55. Y.-J. Du, C.-H. Fu, Explicit BCJ numerators of nonlinear simga model. JHEP **09**, 174 (2016). arXiv:1606.05846 [hep-th]
56. Z. Bern, J.J. Carrasco, L.J. Dixon, H. Johansson, R. Roiban, The ultraviolet behavior of N $=$ 8 supergravity at four loops. Phys. Rev. Lett. **103**, 081301 (2009). arXiv:0905.2326 [hep-th]
57. Z. Bern, S. Davies, T. Dennen, Enhanced ultraviolet cancellations in $\mathcal{N} = 5$ supergravity at four loops. Phys. Rev. D **90**(10), 105011 (2014). arXiv:1409.3089 [hep-th]
58. J.J. Carrasco, H. Johansson, Five-point amplitudes in N $=$ 4 super-Yang–Mills theory and N $=$ 8 supergravity. Phys. Rev. D **85**, 025006 (2012). arXiv:1106.4711 [hep-th]
59. S. Oxburgh, C.D. White, BCJ duality and the double copy in the soft limit. JHEP **02**, 127 (2013). arXiv:1210.1110 [hep-th]
60. Z. Bern, J.J.M. Carrasco, L.J. Dixon, H. Johansson, R. Roiban, Simplifying multiloop integrands and ultraviolet divergences of gauge theory and gravity amplitudes. Phys. Rev. D **85**, 105014 (2012). arXiv:1201.5366 [hep-th]
61. Y.-J. Du, H. Luo, On general BCJ relation at one-loop level in Yang–Mills theory. JHEP **01**, 129 (2013). arXiv:1207.4549 [hep-th]
62. E.Y. Yuan, Virtual color-kinematics duality: 6-pt 1-loop MHV amplitudes. JHEP **05**, 070 (2013). arXiv:1210.1816 [hep-th]
63. R.H. Boels, R.S. Isermann, R. Monteiro, D. O'Connell, Colour-kinematics duality for one-loop rational amplitudes. JHEP **04**, 107 (2013). arXiv:1301.4165 [hep-th]
64. Z. Bern, T. Dennen, Y.-T. Huang, M. Kiermaier, Gravity as the square of gauge theory. Phys. Rev. D **82**, 065003 (2010). arXiv:1004.0693 [hep-th]
65. Z. Bern, J.J.M. Carrasco, L.J. Dixon, H. Johansson, R. Roiban, The complete four-loop four-point amplitude in N $=$ 4 Super–Yang–Mills theory. Phys. Rev. D **82**, 125040 (2010). arXiv:1008.3327 [hep-th]
66. M. Chiodaroli, M. Gunaydin, R. Roiban, Superconformal symmetry and maximal supergravity in various dimensions. JHEP **1203**, 093 (2012). arXiv:1108.3085 [hep-th]
67. Z. Bern, S. Davies, T. Dennen, Y.-T. Huang, Ultraviolet cancellations in half-maximal supergravity as a consequence of the double-copy structure. Phys. Rev. D **86**, 105014 (2012). arXiv:1209.2472 [hep-th]
68. J.J.M. Carrasco, M. Chiodaroli, M. Günaydin, R. Roiban, One-loop four-point amplitudes in pure and matter-coupled N $=$ 4 supergravity. JHEP **1303**, 056 (2013). arXiv:1212.1146 [hep-th]
69. P.H. Damgaard, R. Huang, T. Sondergaard, Y. Zhang, The complete KLT-map between gravity and gauge theories. JHEP **1208**, 101 (2012). arXiv:1206.1577 [hep-th]

70. Y.-T. Huang, H. Johansson, Equivalent D = 3 supergravity amplitudes from double copies of three-algebra and two-algebra gauge theories. Phys. Rev. Lett. **110**, 171601 (2013). arXiv:1210.2255 [hep-th]
71. T. Bargheer, S. He, T. McLoughlin, New relations for three-dimensional supersymmetric scattering amplitudes. Phys. Rev. Lett. **108**, 231601 (2012). arXiv:1203.0562 [hep-th]
72. J.J.M. Carrasco, R. Kallosh, R. Roiban, A.A. Tseytlin, On the U(1) duality anomaly and the S-matrix of N = 4 supergravity. JHEP **07**, 029 (2013). arXiv:1303.6219 [hep-th]
73. M. Chiodaroli, Q. Jin, R. Roiban, Color/kinematics duality for general abelian orbifolds of N = 4 super Yang–Mills theory. JHEP **01**, 152 (2014). arXiv:1311.3600 [hep-th]
74. H. Johansson, A. Ochirov, Pure gravities via color-kinematics duality for fundamental matter. JHEP **11**, 046 (2015). arXiv:1407.4772 [hep-th]
75. M. Chiodaroli, M. Günaydin, H. Johansson, R. Roiban, Scattering amplitudes in $\mathcal{N} = 2$ Maxwell–Einstein and Yang–Mills/Einstein supergravity. JHEP **01**, 081 (2015). arXiv:1408.0764 [hep-th]
76. M. Chiodaroli, M. Gunaydin, H. Johansson, R. Roiban, Spontaneously broken Yang–Mills–Einstein supergravities as double copies. JHEP **06**, 064 (2017). arXiv:1511.01740 [hep-th]
77. M. Chiodaroli, M. Gunaydin, H. Johansson, R. Roiban, Complete construction of magical, symmetric and homogeneous N = 2 supergravities as double copies of gauge theories. Phys. Rev. Lett. **117**(1), 011603 (2016). arXiv:1512.09130 [hep-th]
78. M. Chiodaroli, Simplifying amplitudes in Maxwell–Einstein and Yang–Mills–Einstein supergravities. (2016). arXiv:1607.04129 [hep-th]. https://inspirehep.net/record/1475711/files/arXiv:1607.04129.pdf
79. J.J.M. Carrasco, C.R. Mafra, O. Schlotterer, Semi-abelian Z-theory: NLSM+^{3} from the open string. JHEP **08**, 135 (2017). arXiv:1612.06446 [hep-th]
80. J.J.M. Carrasco, C.R. Mafra, O. Schlotterer, Abelian Z-theory: NLSM amplitudes and α'-corrections from the open string. JHEP **06**, 093 (2017). arXiv:1608.02569 [hep-th]
81. A. Anastasiou, L. Borsten, M.J. Duff, M.J. Hughes, A. Marrani, S. Nagy, M. Zoccali, Twin supergravities from Yang–Mills theory squared. Phys. Rev. D **96**(2), 026013 (2017). arXiv:1610.07192 [hep-th]
82. H. Johansson, G. Kälin, G. Mogull, Two-loop supersymmetric QCD and half-maximal supergravity amplitudes. JHEP **09**, 019 (2017). arXiv:1706.09381 [hep-th]
83. H. Johansson, J. Nohle, Conformal gravity from gauge theory. arXiv:1707.02965 [hep-th]
84. T. Azevedo, O.T. Engelund, Ambitwistor formulations of R^2 gravity and $(DF)^2$ gauge theories. arXiv:1707.02192 [hep-th]
85. A. Anastasiou, L. Borsten, M.J. Duff, A. Marrani, S. Nagy, M. Zoccali, Are all supergravity theories Yang–Mills squared? Nucl. Phys. B **934**, 606–633 (2018). arXiv:1707.03234 [hep-th]
86. M. Chiodaroli, M. Gunaydin, H. Johansson, R. Roiban, Explicit formulae for Yang–Mills–Einstein amplitudes from the double copy. JHEP **07**, 002 (2017). arXiv:1703.00421 [hep-th]
87. M. Chiodaroli, M. Gunaydin, H. Johansson, R. Roiban, Gauged supergravities and spontaneous supersymmetry breaking from the double copy construction. Phys. Rev. Lett. **120**(17), 171601 (2018). arXiv:1710.08796 [hep-th]
88. M. Chiodaroli, M. Günaydin, H. Johansson, R. Roiban, Non-abelian gauged supergravities as double copies. JHEP **06**, 099 (2019). arXiv:1812.10434 [hep-th]
89. T. Azevedo, M. Chiodaroli, H. Johansson, O. Schlotterer, Heterotic and bosonic string amplitudes via field theory. JHEP **10**, 012 (2018). arXiv:1803.05452 [hep-th]
90. Z. Bern, S. Davies, T. Dennen, Y.-T. Huang, Absence of three-loop four-point divergences in N = 4 supergravity. Phys. Rev. Lett. **108**, 201301 (2012). arXiv:1202.3423 [hep-th]
91. Z. Bern, S. Davies, T. Dennen, The ultraviolet structure of half-maximal supergravity with matter multiplets at two and three loops. Phys. Rev. D **88**, 065007 (2013). arXiv:1305.4876 [hep-th]
92. Z. Bern, S. Davies, T. Dennen, A.V. Smirnov, V.A. Smirnov, Ultraviolet properties of N = 4 supergravity at four loops. Phys. Rev. Lett. **111**(23), 231302 (2013). arXiv:1309.2498 [hep-th]
93. Z. Bern, S. Davies, T. Dennen, The ultraviolet critical dimension of half-maximal supergravity at three loops. arXiv:1412.2441 [hep-th]
94. Z. Bern, C. Cheung, H.-H. Chi, S. Davies, L. Dixon, J. Nohle, Evanescent effects can alter ultraviolet divergences in quantum gravity without physical consequences. Phys. Rev. Lett. **115**(21), 211301 (2015). arXiv:1507.06118 [hep-th]
95. Z. Bern, J.J. Carrasco, W.-M. Chen, A. Edison, H. Johansson, J. Parra-Martinez, R. Roiban, M. Zeng, Ultraviolet properties of $\mathcal{N} = 8$ supergravity at five loops. arXiv:1804.09311 [hep-th]

96. S. Deser, J.H. Kay, K.S. Stelle, Renormalizability properties of supergravity. Phys. Rev. Lett. **38**, 527 (1977). arXiv:1506.03757 [hep-th]
97. P.S. Howe, K. Stelle, The ultraviolet properties of supersymmetric field theories. Int. J. Mod. Phys. A **4**, 1871 (1989)
98. M.B. Green, J.G. Russo, P. Vanhove, String theory dualities and supergravity divergences. JHEP **1006**, 075 (2010). arXiv:1002.3805 [hep-th]
99. G. Bossard, P. Howe, K. Stelle, On duality symmetries of supergravity invariants. JHEP **1101**, 020 (2011). arXiv:1009.0743 [hep-th]
100. N. Beisert, H. Elvang, D.Z. Freedman, M. Kiermaier, A. Morales et al., E7(7) constraints on counterterms in N = 8 supergravity. Phys. Lett. B **694**, 265–271 (2010). arXiv:1009.1643 [hep-th]
101. G. Bossard, P. Howe, K. Stelle, P. Vanhove, The vanishing volume of D = 4 superspace. Class. Quant. Gravit. **28**, 215005 (2011). arXiv:1105.6087 [hep-th]
102. G. Bossard, P.S. Howe, K.S. Stelle, Anomalies and divergences in N = 4 supergravity. Phys. Lett. B **719**, 424–429 (2013). arXiv:1212.0841 [hep-th]
103. D.Z. Freedman, R. Kallosh, Y. Yamada, Duality constraints on counterterms in supergravities. Fortsch. Phys. **66**(10), 1800054 (2018). arXiv:1807.06704 [hep-th]
104. Z. Bern, J.J.M. Carrasco, W.-M. Chen, H. Johansson, R. Roiban, M. Zeng, Five-loop four-point integrand of $N = 8$ supergravity as a generalized double copy. Phys. Rev. D **96**(12), 126012 (2017). arXiv:1708.06807 [hep-th]
105. Z. Bern, L.J. Dixon, R. Roiban, Is N = 8 supergravity ultraviolet finite? Phys. Lett. B **644**, 265–271 (2007). arXiv:hep-th/0611086
106. G. Bossard, C. Hillmann, H. Nicolai, E7(7) symmetry in perturbatively quantised N = 8 supergravity. JHEP **12**, 052 (2010). arXiv:1007.5472 [hep-th]
107. R. Kallosh, $E_{7(7)}$ symmetry and finiteness of N = 8 supergravity. JHEP **03**, 083 (2012). arXiv:1103.4115 [hep-th]
108. R. Kallosh, T. Ortin, New E77 invariants and amplitudes. JHEP **1209**, 137 (2012). arXiv:1205.4437 [hep-th]
109. R. Monteiro, D. O'Connell, C.D. White, Black holes and the double copy. JHEP **1412**, 056 (2014). arXiv:1410.0239 [hep-th]
110. A. Luna, R. Monteiro, D. O'Connell, C.D. White, The classical double copy for Taub-NUT spacetime. Phys. Lett. B **750**, 272–277 (2015). arXiv:1507.01869 [hep-th]
111. A.K. Ridgway, M.B. Wise, Static spherically symmetric Kerr–Schild metrics and implications for the classical double copy. Phys. Rev. D **D94**(4), 044023 (2016). arXiv:1512.02243 [hep-th]
112. A. Luna, R. Monteiro, I. Nicholson, D. O'Connell, C.D. White, The double copy: Bremsstrahlung and accelerating black holes. JHEP **06**, 023 (2016). arXiv:1603.05737 [hep-th]
113. C.D. White, Exact solutions for the biadjoint scalar field. Phys. Lett. B **763**, 365–369 (2016). arXiv:1606.04724 [hep-th]
114. W.D. Goldberger, A.K. Ridgway, Radiation and the classical double copy for color charges. Phys. Rev. D **95**(12), 125010 (2017). arXiv:1611.03493 [hep-th]
115. G.L. Cardoso, S. Nagy, S. Nampuri, A double copy for $\mathcal{N} = 2$ supergravity: a linearised tale told on-shell. JHEP **10**, 127 (2016). arXiv:1609.05022 [hep-th]
116. G. Cardoso, S. Nagy, S. Nampuri, Multi-centered $\mathcal{N} = 2$ BPS black holes: a double copy description. JHEP **04**, 037 (2017). arXiv:1611.04409 [hep-th]
117. A. Luna, R. Monteiro, I. Nicholson, A. Ochirov, D. O'Connell, N. Westerberg, C.D. White, Perturbative spacetimes from Yang–Mills theory. JHEP **04**, 069 (2017). arXiv:1611.07508 [hep-th]
118. W.D. Goldberger, S.G. Prabhu, J.O. Thompson, Classical gluon and graviton radiation from the bi-adjoint scalar double copy. Phys. Rev. D **96**(6), 065009 (2017). arXiv:1705.09263 [hep-th]
119. G. Lopes Cardoso, G. Inverso, S. Nagy, S. Nampuri, Comments on the double copy construction for gravitational theories. In 17th Hellenic School and Workshops on Elementary Particle Physics and Gravity (CORFU2017) Corfu, Greece, September 2-28, 2017 (2018). arXiv:1803.07670 [hep-th]. http://inspirehep.net/record/1663475/files/1803.07670.pdf
120. A. Luna, I. Nicholson, D. O'Connell, C.D. White, Inelastic black hole scattering from charged scalar amplitudes. JHEP **03**, 044 (2018). arXiv:1711.03901 [hep-th]
121. N. Bahjat-Abbas, A. Luna, C.D. White, The Kerr–Schild double copy in curved spacetime. JHEP **12**, 004 (2017). arXiv:1710.01953 [hep-th]
122. D.S. Berman, E. Chacón, A. Luna, C.D. White, The self-dual classical double copy, and the Eguchi–Hanson instanton. JHEP **01**, 107 (2019). arXiv:1809.04063 [hep-th]

123. J. Plefka, J. Steinhoff, W. Wormsbecher, Effective action of dilaton gravity as the classical double copy of Yang–Mills theory. Phys. Rev. D **99**(2), 024021 (2019). arXiv:1807.09859 [hep-th]
124. N. Bahjat-Abbas, R. Stark-Muchão, C.D. White, Biadjoint wires. Phys. Lett. B **788**, 274–279 (2019). arXiv:1810.08118 [hep-th]
125. A. Luna, R. Monteiro, I. Nicholson, D. O'Connell, Type D spacetimes and the weyl double copy. Class. Quant. Gravit. **36**, 065003 (2019). arXiv:1810.08183 [hep-th]
126. C.-H. Shen, Gravitational radiation from Color-kinematics duality. JHEP **11**, 162 (2018). arXiv:1806.07388 [hep-th]
127. C. Cheung, I.Z. Rothstein, M.P. Solon, From scattering amplitudes to classical potentials in the post-Minkowskian expansion. Phys. Rev. Lett. **121**(25), 251101 (2018). arXiv:1808.02489 [hep-th]
128. D.A. Kosower, B. Maybee, D. O'Connell, Amplitudes, observables, and classical scattering. JHEP **02**, 137 (2019). arXiv:1811.10950 [hep-th]
129. M. Carrillo González, B. Melcher, K. Ratliff, S. Watson, C.D. White, The classical double copy in three spacetime dimensions. JHEP **07**, 167 (2019). arXiv:1904.11001 [hep-th]
130. H. Johansson, A. Ochirov, Double copy for massive quantum particles with spin. JHEP **09**, 040 (2019). arXiv:1906.12292 [hep-th]
131. B. Maybee, D. O'Connell, J. Vines, Observables and amplitudes for spinning particles and black holes. arXiv:1906.09260 [hep-th]
132. J. Plefka, C. Shi, J. Steinhoff, T. Wang, Breakdown of the classical double copy for the effective action of dilaton-gravity at NNLO. Phys. Rev. D **100**(8), 086006 (2019). arXiv:1906.05875 [hep-th]
133. Z. Bern, C. Cheung, R. Roiban, C.-H. Shen, M.P. Solon, M. Zeng, Scattering amplitudes and the conservative hamiltonian for binary systems at third post-minkowskian order. Phys. Rev. Lett. **122**(20), 201603 (2019). arXiv:1901.04424 [hep-th]
134. Z. Bern, C. Cheung, R. Roiban, C.-H. Shen, M.P. Solon, M. Zeng, Black hole binary dynamics from the double copy and effective theory. arXiv:1908.01493 [hep-th]
135. N. Arkani-Hamed, Y.-T. Huang, D. O'Connell, Kerr black holes as elementary particles. arXiv:1906.10100 [hep-th]
136. L. Borsten, M.J. Duff, L.J. Hughes, S. Nagy, A magic square from Yang–Mills squared. Phys. Rev. Lett. **112**, 131601 (2014). arXiv:1301.4176 [hep-th]
137. A. Anastasiou, L. Borsten, M.J. Duff, L.J. Hughes, S. Nagy, Yang–Mills origin of gravitational symmetries. Phys. Rev. Lett. **113**(23), 231606 (2014). arXiv:1408.4434 [hep-th]
138. A. Anastasiou, L. Borsten, M.J. Duff, L.J. Hughes, S. Nagy, A magic pyramid of supergravities. JHEP **1404**, 178 (2014). arXiv:1312.6523 [hep-th]
139. L. Borsten, M.J. Duff, Gravity as the square of Yang–Mills? Phys. Script. **90**, 108012 (2015). arXiv:1602.08267 [hep-th]
140. A. Anastasiou, L. Borsten, L.J. Hughes, S. Nagy, Global symmetries of Yang–Mills squared in various dimensions. JHEP **148**, 1601 (2016). arXiv:1502.05359 [hep-th]
141. A. Anastasiou, L. Borsten, M.J. Duff, S. Nagy, M. Zoccali, Gravity as gauge theory squared: a ghost story. Phys. Rev. Lett. **121**(21), 211601 (2018). arXiv:1807.02486 [hep-th]
142. L. Borsten, $D = 6$, $\mathcal{N} = (2, 0)$ and $\mathcal{N} = (4, 0)$ theories. Phys. Rev. D **97**, 066014 (2018). arXiv:1708.02573 [hep-th]
143. J.J.M. Carrasco, Gauge and Gravity Amplitude Relations. In Journeys Through the Precision Frontier: Amplitudes for Colliders. World Scientific, Singapore (2015). arXiv:1506.00974 [hep-th]
144. Z. Bern, J.J. Carrasco, M. Chiodaroli, H. Johansson, R. Roiban, The duality between color and kinematics and its applications. arXiv:1909.01358 [hep-th]
145. M.L. Mangano, S.J. Parke, Multiparton amplitudes in gauge theories. Phys. Rept. **200**, 301–367 (1991). arXiv:hep-th/0509223 [hep-th]
146. J.M. Henn, J.C. Plefka, Scattering amplitudes in gauge theories. Lect. Notes Phys. **883**, 1–195 (2014)
147. C.D. White, Aspects of high energy scattering. arXiv:1909.05177 [hep-th]
148. C. Cheung, TASI lectures on scattering amplitudes. In Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating the Next Discoveries in Particle Physics (TASI 2016): Boulder, CO, USA, June 6–July 1, 2016, pp. 571–623. (2018). arXiv:1708.03872 [hep-ph]
149. Z. Bern, Perturbative quantum gravity and its relation to gauge theory. Living Rev. Rel. **5**, 5 (2002). arXiv:gr-qc/0206071 [gr-qc]
150. T. Adamo, Lectures on twistor theory. PoS **Modave2017**, 003 (2018). arXiv:1712.02196 [hep-th]
151. Y. Geyer, Ambitwistor Strings: Worldsheet Approaches to perturbative Quantum Field Theories. PhD thesis, Oxford University, Inst. Math. (2016). arXiv:1610.04525 [hep-th]

152. L.D. Faddeev, V.N. Popov, Feynman diagrams for the Yang–Mills field. Phys. Lett. B **25**, 29–30 (1967). [325(1967)]
153. C. Becchi, A. Rouet, R. Stora, Renormalization of gauge theories. Ann. Phys. **98**, 287–321 (1976)
154. I.V. Tyutin, Gauge invariance in field theory and statistical physics in operator formalism. Lebedev Physics Institute preprint **39** (1975). arXiv:0812.0580 [hep-th]
155. T. Kugo, I. Ojima, Local covariant operator formalism of nonabelian gauge theories and quark confinement problem. Prog. Theor. Phys. Suppl. **66**, 1–130 (1979)
156. M. Henneaux, C. Teitelboim, Quantization of gauge systems. (1992)
157. J. Gomis, J. Paris, S. Samuel, Antibracket, antifields and gauge theory quantization. Phys. Rept. **259**, 1–145 (1995). arXiv:hep-th/9412228 [hep-th]
158. E.S. Fradkin, G.A. Vilkovisky, Quantization of relativistic systems with constraints. Phys. Lett. B **55**, 224–226 (1975)
159. I.A. Batalin, G.A. Vilkovisky, Relativistic S matrix of dynamical systems with Boson and Fermion constraints. Phys. Lett. B **69B**, 309–312 (1977)
160. I.A. Batalin, G.A. Vilkovisky, Gauge algebra and quantization. Phys. Lett. B **102**, 27–31 (1981)
161. I.A. Batalin, G.A. Vilkovisky, Quantization of gauge theories with linearly dependent generators. Phys. Rev. D **28**, 2567–2582 (1983). [Erratum: Phys. Rev.D30,508(1984)]
162. I.A. Batalin, G.A. Vilkovisky, Closure of the gauge algebra, generalized lie equations and Feynman rules. Nucl. Phys. B **234**, 106–124 (1984)
163. I.A. Batalin, G.A. Vilkovisky, Existence theorem for gauge algebra. J. Math. Phys. **26**, 172–184 (1985)
164. M. Srednicki, *Quantum Field Theory* (Cambridge University Press, Cambridge, 2007)
165. D.-P. Zhu, Zeros in Scattering amplitudes and the structure of nonabelian gauge theories. Phys. Rev. D **22**, 2266 (1980)
166. C.J. Goebel, F. Halzen, J.P. Leveille, Angular zeros of Brown, Mikaelian, Sahdev, and Samuel and the factorization of tree amplitudes in gauge theories. Phys. Rev. D **23**, 2682–2685 (1981)
167. R. Kleiss, H. Kuijf, Multi-gluon cross-sections and five jet production at hadron colliders. Nucl. Phys. B **312**, 616–644 (1989)
168. V. Del Duca, L.J. Dixon, F. Maltoni, New color decompositions for gauge amplitudes at tree and loop level. Nucl. Phys. B **571**, 51–70 (2000). arXiv:hep-ph/9910563 [hep-ph]
169. N.E.J. Bjerrum-Bohr, P.H. Damgaard, P. Vanhove, Minimal basis for gauge theory amplitudes. Phys. Rev. Lett. **103**, 161602 (2009). arXiv:0907.1425 [hep-th]
170. S. Stieberger, Open and closed vs. pure open string disk amplitudes. arXiv:0907.2211 [hep-th]
171. C.R. Mafra, O. Schlotterer, Berends–Giele recursions and the BCJ duality in superspace and components. JHEP **03**, 097 (2016). arXiv:1510.08846 [hep-th]
172. S.H. Henry Tye, Y. Zhang, Dual identities inside the gluon and the graviton scattering amplitudes. JHEP **06**, 071 (2010). arXiv:1003.1732 [hep-th]. [Erratum: JHEP04,114(2011)]
173. C.R. Mafra, O. Schlotterer, The structure of n-point one-loop open superstring amplitudes. JHEP **08**, 099 (2014). arXiv:1203.6215 [hep-th]
174. O. Schlotterer, S. Stieberger, Motivic multiple zeta values and superstring amplitudes. J. Phys. **A46**, 475401 (2013). arXiv:1205.1516 [hep-th]
175. J. Broedel, O. Schlotterer, S. Stieberger, Polylogarithms, multiple zeta values and superstring amplitudes. Fortsch. Phys. **61**, 812–870 (2013). arXiv:1304.7267 [hep-th]
176. B. Feng, R. Huang, Y. Jia, Gauge amplitude identities by on-shell recursion relation in S-matrix program. Phys. Lett. B **B695**, 350–353 (2011). arXiv:1004.3417 [hep-th]
177. R. Britto, F. Cachazo, B. Feng, E. Witten, Direct proof of tree-level recursion relation in Yang–Mills theory. Phys. Rev. Lett. **94**, 181602 (2005). arXiv:hep-th/0501052 [hep-th]
178. F. Cachazo, Fundamental BCJ relation in $N = 4$ SYM from the connected formulation. arXiv:1206.5970 [hep-th]
179. D. Vaman, Y.-P. Yao, Constraints and generalized gauge transformations on tree-level gluon and graviton amplitudes. JHEP **11**, 028 (2010). arXiv:1007.3475 [hep-th]
180. A. Ferber, Supertwistors and conformal supersymmetry. Nucl. Phys. B **132**, 55–64 (1978)
181. Z. Bern, S. Davies, T. Dennen, Y.-T. Huang, J. Nohle, Color-kinematics duality for pure Yang–Mills and gravity at one and two loops. Phys. Rev. D **92**(4), 045041 (2015). arXiv:1303.6605 [hep-th]
182. J.J.M. Carrasco, H. Johansson, Generic multiloop methods and application to $N = 4$ super–Yang–Mills. J. Phys. A **44**, 454004 (2011). arXiv:1103.3298 [hep-th]

183. Z. Bern, C. Boucher-Veronneau, H. Johansson, N> = 4 Supergravity amplitudes from gauge theory at one loop. Phys. Rev. D **84**, 105035 (2011). arXiv:1107.1935 [hep-th]
184. C. Boucher-Veronneau, L.J. Dixon, N> − 4 supergravity amplitudes from gauge theory at two loops. JHEP **12**, 046 (2011). arXiv:1110.1132 [hep-th]
185. C.R. Mafra, O. Schlotterer, Two-loop five-point amplitudes of super Yang–Mills and supergravity in pure spinor superspace. JHEP **10**, 124 (2015). arXiv:1505.02746 [hep-th]
186. S. Weinzierl, Fermions and the scattering equations. JHEP **03**, 141 (2015). arXiv:1412.5993 [hep-th]
187. T. Kugo, P.K. Townsend, Supersymmetry and the division algebras. Nucl. Phys. B **221**, 357 (1983)
188. J.M. Evans, Supersymmetric Yang–Mills theories and division algebras. Nucl. Phys. B **298**, 92 (1988)
189. J.C. Baez, J. Huerta, Division algebras and supersymmetry I. In R. Doran, G. Friedman, J. Rosenberg, (eds.), Superstrings, Geometry, Topology, and C*-Algebras, Proc. Symp. Pure Math, vol. 81, 65–80. (2009). arXiv:0909.0551 [hep-th]
190. A. Anastasiou, L. Borsten, M.J. Duff, L.J. Hughes, S. Nagy, Super Yang–Mills, division algebras and triality. JHEP **1408**, 080 (2014). arXiv:1309.0546 [hep-th]
191. A. Sudbery, Division algebras, (pseudo)orthogonal groups, and spinors. J. Phys. A **17**(5), 939–955 (1984)
192. K. Chung, A. Sudbery, Octonions and the Lorentz and conformal groups of ten-dimensional space-time. Phys. Lett. B **198**, 161 (1987)
193. C.A. Manogue, A. Sudbery, General solutions of covariant superstring equations of motion. Phys. Rev. D **40**, 4073 (1989)
194. P. Ramond, Introduction to exceptional lie groups and algebras. CALT-68-577 (1977)
195. J.C. Baez, The octonions. Bull. Am. Math. Soc. **39**, 145–205 (2002). arXiv:math/0105155 [math-ra]
196. C.H. Barton, A. Sudbery, Magic squares and matrix models of Lie algebras. Adv. Math. **180**(2), 596–647 (2003). arXiv:math/0203010
197. A. Anastasiou, L. Borsten, M.J. Duff, A. Marrani, S. Nagy, M. Zoccali, The mile high magic pyramid. In P. Vojtechovsky, (ed.), Nonassociative mathematics and its applications, vol. 721 of Contemporary Mathematics, pp. 1–27. Amer. Math. Soc., Providence, RI (2019). arXiv:1711.08476 [hep-th]
198. A. Hodges, New expressions for gravitational scattering amplitudes. J. High Energy Phys. **1307** (2013). arXiv:1108.2227 [hep-th]
199. F. Cachazo, S. He, E.Y. Yuan, Scattering of massless particles: scalars, gluons and gravitons. JHEP **1407**, 033 (2014). arXiv:1309.0885 [hep-th]
200. R. Monteiro, D. O'Connell, The kinematic algebras from the scattering equations. JHEP **1403**, 110 (2014). arXiv:1311.1151 [hep-th]
201. F. Cachazo, S. He, E.Y. Yuan, Scattering equations and matrices: from Einstein to Yang–Mills, DBI and NLSM. JHEP **07**, 149 (2015). arXiv:1412.3479 [hep-th]
202. S.G. Naculich, Scattering equations and BCJ relations for gauge and gravitational amplitudes with massive scalar particles. JHEP **09**, 029 (2014). arXiv:1407.7836 [hep-th]
203. S.G. Naculich, CHY representations for gauge theory and gravity amplitudes with up to three massive particles. JHEP **05**, 050 (2015). arXiv:1501.03500 [hep-th]
204. C. Cheung, C.-H. Shen, Symmetry for flavor-kinematics duality from an action. Phys. Rev. Lett. **118**(12), 121601 (2017). arXiv:1612.00868 [hep-th]
205. R.W. Brown, S.G. Naculich, KLT-type relations for QCD and bicolor amplitudes from color-factor symmetry. JHEP **03**, 057 (2018). arXiv:1802.01620 [hep-th]
206. H. Johansson, A. Ochirov, Color-kinematics duality for QCD amplitudes. JHEP **01**, 170 (2016). arXiv:1507.00332 [hep-ph]
207. M. Chiodaroli, Simplifying amplitudes in Maxwell–Einstein and Yang–Mills–Einstein supergravities. In J. Brüning, M. Staudacher, (eds) Space-time–matter: analytic and geometric structures (2018)
208. N. Marcus, J.H. Schwarz, Three-dimensional supergravity theories. Nucl. Phys. B **228**, 145 (1983)
209. B. de Wit, A. Tollsten, H. Nicolai, Locally supersymmetric d = 3 nonlinear sigma models. Nucl. Phys. B **392**, 3–38. arXiv:hep-th/9208074
210. J. Bagger, N. Lambert, Modeling multiple M2's. Phys. Rev. D **75**, 045020 (2007). arXiv:hep-th/0611108 [hep-th]
211. A. Gustavsson, Algebraic structures on parallel M2-branes. Nucl. Phys. B **811**, 66–76 (2009). arXiv:0709.1260 [hep-th]
212. J. Bagger, N. Lambert, Gauge symmetry and supersymmetry of multiple M2-branes. Phys. Rev. D **77**, 065008 (2008). arXiv:0711.0955 [hep-th]

213. Y.-T. Huang, H. Johansson, S. Lee, On three-algebra and bi-fundamental matter amplitudes and integrability of supergravity. JHEP **11**, 050 (2013). arXiv:1307.2222 [hep-th]
214. O. Aharony, O. Bergman, D.L. Jafferis, J. Maldacena, N = 6 superconformal Chern–Simons-matter theories, M2-branes and their gravity duals. JHEP **10**, 091 (2008). arXiv:0806.1218 [hep-th]
215. S. Palmer, C. Sämann, The ABJM model is a higher gauge theory. Int. J. Geom. Methods. Mod. Phys. **11**(08), 1450075 (2014). arXiv:1311.1997 [hep-th]
216. A. Ochirov, P. Tourkine, BCJ duality and double copy in the closed string sector. JHEP **05**, 136 (2014). arXiv:1312.1326 [hep-th]
217. N.E.J. Bjerrum-Bohr, P.H. Damgaard, P. Tourkine, P. Vanhove, Scattering equations and string theory amplitudes. Phys. Rev. D **90**(10), 106002 (2014). arXiv:1403.4553 [hep-th]
218. S. He, R. Monteiro, O. Schlotterer, String-inspired BCJ numerators for one-loop MHV amplitudes. JHEP **01**, 171 (2016). arXiv:1507.06288 [hep-th]
219. L. Dolan, P. Goddard, Proof of the formula of Cachazo, He and Yuan for Yang–Mills tree amplitudes in arbitrary dimension. JHEP **1405**, 010 (2014). arXiv:1311.5200 [hep-th]
220. N.E.J. Bjerrum-Bohr, J.L. Bourjaily, P.H. Damgaard, B. Feng, Manifesting Color–kinematics duality in the scattering equation formalism. JHEP **09**, 094 (2016). arXiv:1608.00006 [hep-th]
221. L. Mason, D. Skinner, Ambitwistor strings and the scattering equations. JHEP **1407**, 048 (2014). arXiv:1311.2564 [hep-th]
222. T. Adamo, E. Casali, D. Skinner, Ambitwistor strings and the scattering equations at one loop. JHEP **04**, 104 (2014). arXiv:1312.3828 [hep-th]
223. Y. Geyer, L. Mason, R. Monteiro, P. Tourkine, Loop integrands for scattering amplitudes from the Riemann sphere. Phys. Rev. Lett. **115**(12), 121603 (2015). arXiv:1507.00321 [hep-th]
224. Y. Geyer, L. Mason, R. Monteiro, P. Tourkine, Two-loop scattering amplitudes from the Riemann sphere. arXiv:1607.08887 [hep-th]
225. T. Adamo, E. Casali, L. Mason, S. Nekovar, Scattering on plane waves and the double copy. Class. Quant. Gravit. **35**(1), 015004 (2018). arXiv:1706.08925 [hep-th]
226. M.J. Duff, P. van Nieuwenhuizen, Quantum inequivalence of different field representations. Phys. Lett. B **94**, 179 (1980)
227. R. Gilmore, *Lie Groups, Lie Algebras, and Some of Their Applications* (Courier Corporation, Massachusetts, 2012)
228. A.L. Besse, Einstein Manifolds. A Series of Modern Surveys in Mathematics. Springer, Berlin (1987). http://link.springer.com/book/10.1007/BF02345020
229. M. Ben-Shahar, M. Chiodaroli, One-loop amplitudes for $\mathcal{N} = 2$ homogeneous supergravities. JHEP **03**, 153 (2019). arXiv:1812.00402 [hep-th]
230. B.S. DeWitt, Quantum theory of gravity. 1. The canonical theory. Phys. Rev. **160**, 1113–1148 (1967). [**3**, 93 (1987)]
231. B. S. DeWitt, Quantum theory of gravity. 2. The manifestly covariant theory. Phys. Rev. **162**, 1195–1239 (1967). [298 (1967)]
232. B. S. DeWitt, Quantum theory of gravity. 3. Applications of the covariant theory. Phys. Rev. **162**, 1239–1256 (1967). [307 (1967)]
233. N.E.J. Bjerrum-Bohr, P.H. Damgaard, B. Feng, T. Sondergaard, Gravity and Yang–Mills amplitude relations. Phys. Rev. D **82**, 107702 (2010). arXiv:1005.4367 [hep-th]
234. L. Castellani, A. Ceresole, S. Ferrara, R. D'Auria, P. Fre et al., The complete n = 3 matter coupled supergravity. Nucl. Phys. B **268**, 317 (1986)
235. A. Strominger, Special geometry. Commun. Math. Phys. **133**, 163–180 (1990)
236. B. de Wit, A. Van Proeyen, Broken sigma model isometries in very special geometry. Phys. Lett. B **293**, 94–99 (1992). arXiv:hep-th/9207091 [hep-th]
237. M. Günaydin, G. Sierra, P.K. Townsend, The geometry of $N = 2$ Maxwell–Einstein supergravity and Jordan algebras. Nucl. Phys. B **242**, 244 (1984)
238. B. de Wit, A. Van Proeyen, Potentials and symmetries of general gauged N = 2 supergravity: Yang–Mills models. Nucl. Phys. B **B245**, 89–117 (1984)
239. J. Bagger, E. Witten, Matter couplings in N = 2 supergravity. Nucl. Phys. B **222**, 1–10 (1983)
240. B. de Wit, P.G. Lauwers, R. Philippe, S.Q. Su, A. Van Proeyen, Gauge and matter fields coupled to N = 2 supergravity. Phys. Lett. B **134**, 37–43 (1984)
241. S. Cecotti, S. Ferrara, L. Girardello, Geometry of type II superstrings and the moduli of superconformal field theories. Int. J. Mod. Phys. A **4**, 2475 (1989)

242. S. Cecotti, Homogeneous Kahler manifolds and T algebras in $N = 2$ supergravity and superstrings. Commun. Math. Phys. **124**, 23–55 (1989)
243. B. de Wit, F. Vanderseypen, A. Van Proeyen, Symmetry structure of special geometries. Nucl. Phys. B **400**, 463–524 (1993). arXiv:hep-th/9210068 [hep-th]
244. M. Günaydin, G. Sierra, P.K. Townsend, Gauging the $d = 5$ Maxwell–Einstein supergravity theories: more on Jordan algebras. Nucl. Phys. B **B253**, 573 (1985)
245. M. Günaydin, G. Sierra, P.K. Townsend, Exceptional supergravity theories and the magic square. Phys. Lett. B **133**, 72 (1983)
246. B. de Wit, A. Van Proeyen, Isometries of special manifolds (1995). arXiv:hep-th/9505097 [hep-th]
247. K. Saraikin, C. Vafa, Non-supersymmetric black holes and topological strings. Class. Quant. Gravit. **25**, 095007 (2008). arXiv:hep-th/0703214 [hep-th]
248. M.J. Duff, J.T. Liu, J. Rahmfeld, Four-dimensional string–string–string triality. Nucl. Phys. B **459**, 125–159 (1996). arXiv:hep-th/9508094
249. D. Nandan, J. Plefka, O. Schlotterer, C. Wen, Einstein–Yang–Mills from pure Yang–Mills amplitudes. arXiv:1607.05701 [hep-th]
250. F. Teng, B. Feng, Expanding Einstein–Yang–Mills by Yang–Mills in CHY frame. JHEP **05**, 075 (2017). arXiv:1703.01269 [hep-th]
251. Y.-J. Du, B. Feng, F. Teng, Expansion of all multitrace tree level EYM amplitudes. JHEP **12**, 038 (2017). arXiv:1708.04514 [hep-th]
252. S. He, O. Schlotterer, New relations for gauge-theory and gravity amplitudes at loop level. Phys. Rev. Lett. **118**(16), 161601 (2017). arXiv:1612.00417 [hep-th]
253. D. Nandan, J. Plefka, G. Travaglini, All rational one-loop Einstein–Yang–Mills amplitudes at four points. JHEP **09**, 011 (2018). arXiv:1803.08497 [hep-th]
254. Y. Dolivet, B. Julia, C. Kounnas, Magic N = 2 supergravities from hyper-free superstrings. JHEP **02**, 097 (2008). arXiv:0712.2867 [hep-th]
255. M. Bianchi, S. Ferrara, Enriques and octonionic magic supergravity models. JHEP **02**, 054 (2008). arXiv:0712.2976 [hep-th]
256. D. Roest, H. Samtleben, Twin supergravities. Class. Quant. Gravit. **26**, 155001 (2009). arXiv:0904.1344 [hep-th]
257. M.J. Duff, S. Ferrara, Generalized mirror symmetry and trace anomalies. Class. Quant. Gravit. **28**, 065005 (2011). arXiv:1009.4439 [hep-th]
258. L. Borsten, M.J. Duff, A. Marrani, Twin conformal field theories. JHEP **03**, 112 (2019). arXiv:1812.11130 [hep-th]
259. H. Johansson, G. Mogull, F. Teng, Unraveling conformal gravity amplitudes. JHEP **09**, 080 (2018). arXiv:1806.05124 [hep-th]
260. N. Berkovits, E. Witten, Conformal supergravity in twistor-string theory. JHEP **08**, 009 (2004). arXiv:hep-th/0406051 [hep-th]
261. E.S. Fradkin, A.A. Tseytlin, Conformal supergravity. Phys. Rept. **119**, 233–362 (1985)
262. C.R. Mafra, O. Schlotterer, Non-abelian Z-theory: Berends–Giele recursion for the α'-expansion of disk integrals. JHEP **01**, 031 (2017). arXiv:1609.07078 [hep-th]
263. O. Schlotterer, O. Schnetz, Closed strings as single-valued open strings: a genus-zero derivation. J. Phys. A **52**(4), 045401 (2019). arXiv:1808.00713 [hep-th]
264. F. Cachazo, P. Cha, S. Mizera, Extensions of theories from soft limits. JHEP **06**, 170 (2016). arXiv:1604.03893 [hep-th]
265. J.A. Farrow, A.E. Lipstein, P. McFadden, Double copy structure of CFT correlators. JHEP **02**, 130 (2019). arXiv:1812.11129 [hep-th]
266. A.R. Fazio, Cosmological correlators. In: In Formalism and double copy. arXiv:1909.07343 [hep-th]
267. A. Lipstein, P. McFadden, Double copy structure and the flat space limit of conformal correlators in even dimensions. arXiv:1912.10046 [hep-th]
268. S. Ferrara, D. Lüst, Spin-four $\mathcal{N} = 7$ W-supergravity: S-fold and double copy construction. JHEP **07**, 114 (2018). arXiv:1805.10022 [hep-th]
269. Y. Geyer, L. Mason, Polarized scattering equations for 6D superamplitudes. Phys. Rev. Lett. **122**(10), 101601 (2019). arXiv:1812.05548 [hep-th]
270. S. Ferrara, M. Porrati, A. Zaffaroni, N = 6 supergravity on AdS(5) and the SU(2,2/3) superconformal correspondence. Lett. Math. Phys. **47**, 255–263 (1999). arXiv:hep-th/9810063 [hep-th]
271. O. Aharony, M. Evtikhiev, On four dimensional N = 3 superconformal theories. JHEP **04**, 040 (2016). arXiv:1512.03524 [hep-th]

272. I. García-Etxebarria, D. Regalado, $\mathcal{N} = 3$ four dimensional field theories. JHEP **03**, 083 (2016). arXiv:1512.06434 [hep-th]
273. O. Aharony, Y. Tachikawa, S-folds and 4d N = 3 superconformal field theories. JHEP **06**, 044 (2016). arXiv:1602.08638 [hep-th]
274. Y. Geyer, A.E. Lipstein, L.J. Mason, Ambitwistor strings in four dimensions. Phys. Rev. Lett. **113**(8), 081602 (2014). arXiv:1404.6219 [hep-th]
275. Y. Geyer, L. Mason, R. Monteiro, P. Tourkine, One-loop amplitudes on the Riemann sphere. JHEP **03**, 114 (2016). arXiv:1511.06315 [hep-th]
276. C. Hull, Strongly coupled gravity and duality. Nucl. Phys. B **583**, 237–259 (2000). arXiv:hep-th/0004195 [hep-th]
277. R. Monteiro, D. O'Connell, The kinematic algebra from the self-dual sector. JHEP **1107**, 007 (2011). arXiv:1105.2565 [hep-th]
278. N. Bjerrum-Bohr, P.H. Damgaard, R. Monteiro, D. O'Connell, Algebras for amplitudes. JHEP **1206**, 061 (2012). arXiv:1203.0944 [hep-th]
279. M. Tolotti, S. Weinzierl, Construction of an effective Yang–Mills Lagrangian with manifest BCJ duality. JHEP **07**, 111 (2013). arXiv:1306.2975 [hep-th]
280. C.-H. Fu, K. Krasnov, Colour-Kinematics duality and the Drinfeld double of the Lie algebra of diffeomorphisms. JHEP **01**, 075 (2017). arXiv:1603.02033 [hep-th]
281. C. Cheung, G.N. Remmen, Twofold symmetries of the pure gravity action. JHEP **01**, 104 (2017). arXiv:1612.03927 [hep-th]
282. G. Chen, H. Johansson, F. Teng, T. Wang, On the kinematic algebra for BCJ numerators beyond the MHV sector. arXiv:1906.10683 [hep-th]
283. Z. Bern, A.K. Grant, Perturbative gravity from QCD amplitudes. Phys. Lett. B **457**, 23–32 (1999). arXiv:hep-th/9904026 [hep-th]
284. O. Hohm, On factorizations in perturbative quantum gravity. JHEP **04**, 103 (2011). arXiv:1103.0032 [hep-th]
285. C. Cheung, G.N. Remmen, Hidden simplicity of the gravity action. JHEP **09**, 002 (2017). arXiv:1705.00626 [hep-th]
286. W. Siegel, Two vierbein formalism for string inspired axionic gravity. Phys. Rev. D **47**, 5453–5459 (1993). arXiv:hep-th/9302036 [hep-th]
287. W. Siegel, Superspace duality in low-energy superstrings. Phys. Rev. D **48**, 2826–2837 (1993). arXiv:hep-th/9305073 [hep-th]
288. C. Hull, B. Zwiebach, Double field theory. JHEP **0909**, 099 (2009). arXiv:0904.4664 [hep-th]
289. C. Hull, B. Zwiebach, The Gauge algebra of double field theory and Courant brackets. JHEP **0909**, 090 (2009). arXiv:0908.1792 [hep-th]
290. O. Hohm, C. Hull, B. Zwiebach, Generalized metric formulation of double field theory. JHEP **1008**, 008 (2010). arXiv:1006.4823 [hep-th]
291. M.J. Duff, Duality rotations in string theory. Nucl. Phys. B **335**, 610 (1990)
292. M.J. Duff, J.X. Lu, Duality rotations in membrane theory. Nucl. Phys. B **347**, 394–419 (1990)
293. A.A. Tseytlin, Duality symmetric formulation of string world sheet dynamics. Phys. Lett. B **242**, 163–174 (1990)
294. A.A. Tseytlin, Duality symmetric closed string theory and interacting chiral scalars. Nucl. Phys. B **350**, 395–440 (1991)
295. A.I. Janis, E.T. Newman, J. Winicour, Reality of the Schwarzschild singularity. Phys. Rev. Lett. **20**, 878–880 (1968)
296. K. Kim, K. Lee, R. Monteiro, I. Nicholson, D. Peinador Veiga, The classical double copy of a point charge. arXiv:1912.02177 [hep-th]
297. M.J. Duff, J. Lu, Duality for strings and membranes. Conf. Proc. C **9003122**, 148–162 (1990)
298. K. Lee, Kerr–Schild double field theory and classical double copy. JHEP **10**, 027 (2018). arXiv:1807.08443 [hep-th]
299. W. Siegel, Superstrings give old minimal supergravity. Phys. Lett. B **211**, 55 (1988)
300. W. Siegel, Curved extended superspace from Yang–Mills theory a la strings. Phys. Rev. D **53**, 3324–3336 (1996). arXiv:hep-th/9510150 [hep-th]
301. C.M. Hull, The BRST and anti-BRST invariant quantization of general gauge theories. Mod. Phys. Lett. **A5**, 1871–1882 (1990)
302. L. Borsten, S. Nagy, The classical double-copy: ghosts and gauges. In preparation (2019)
303. M. Zoccali, Supergravity as Yang–Mills squared. PhD thesis, Imperial Coll., London (2018)

304. L. Borsten, I. Jubb, V. Makwana, S. Nagy, Gauge × gauge on spheres. arXiv:1911.12324 [hep-th]
305. S. Cecotti, S. Ferrara, M. Porrati, S. Sabharwal, New minimal higher derivative supergravity coupled to matter. Nucl. Phys. B **306**, 160 (1988)
306. S. Ferrara, S. Sabharwal, Structure of new minimal supergravity. Ann. Phys. **189**, 318–351 (1989)
307. K. Stelle, P.C. West, Minimal auxiliary fields for supergravity. Phys. Lett. B **74**, 330 (1978)
308. S. Ferrara, P. van Nieuwenhuizen, The auxiliary fields of supergravity. Phys. Lett. B **74**, 333 (1978)
309. E. Cremmer, B. Julia, The $SO(8)$ supergravity. Nucl. Phys. B **159**, 141 (1979)
310. C.M. Hull, P.K. Townsend, Unity of superstring dualities. Nucl. Phys. B **438**, 109–137 (1995). arXiv:hep-th/9410167
311. M. Bianchi, H. Elvang, D.Z. Freedman, Generating tree amplitudes in N = 4 SYM and N = 8 SG. JHEP **0809**, 063 (2008). arXiv:0805.0757 [hep-th]
312. N. Arkani-Hamed, F. Cachazo, J. Kaplan, What is the simplest quantum field theory? JHEP **09**, 016 (2010). arXiv:0808.1446 [hep-th]
313. B. Julia, Group disintegrations. In: S. Hawking, M. Rocek, (eds.), Superspace and Supergravity, vol. C8006162 of Nuffield Gravity Workshop, Cambridge University Press, Cambridge, pp. 331–350 (1980)
314. H. Nicolai, The integrability of $N = 16$ supergravity. Phys. Lett. B **194**, 402 (1987)
315. P.C. West, E(11) and M theory. Class. Quant. Gravit. **18**, 4443–4460 (2001). arXiv:hep-th/0104081 [hep-th]
316. T. Damour, M. Henneaux, H. Nicolai, E(10) and a 'small tension expansion' of M theory. Phys. Rev. Lett. **89**, 221601 (2002). arXiv:hep-th/0207267 [hep-th]
317. H. Freudenthal, Beziehungen der E_7 und E_8 zur oktavenebene I-II. Nederl. Akad. Wetensch. Proc. Ser. **57**, 218–230 (1954)
318. J. Tits, Interprétation géométriques de groupes de Lie simples compacts de la classe E. Mém. Acad. R. Belg. Sci **29**, 3 (1955)
319. B.A. Rosenfeld, Geometrical interpretation of the compact simple Lie groups of the class E (Russian). Dokl. Akad. Nauk SSSR **106**, 600–603 (1956)
320. S. L. Adler, Consistency conditions on the strong interactions implied by a partially conserved axial vector current. Phys. Rev. **137**, B1022–B1033 (1965) [140 (1964)]
321. N. Craig, H. Elvang, M. Kiermaier, T. Slatyer, Massive amplitudes on the Coulomb branch of N = 4 SYM. JHEP **12**, 097 (2011). arXiv:1104.2050 [hep-th]
322. H. Freudenthal, Beziehungen der E_7 und E_8 zur oktavenebene IX. Nederl. Akad. Wetensch. Proc. Ser. **A62**, 466–474 (1959)
323. J. Tits, Algébres alternatives, algébres de jordan et algébres de lie exceptionnelles. Indag. Math. **28**, 223–237 (1966)
324. S.L. Cacciatori, B.L. Cerchiai, A. Marrani, Squaring the magic. Adv. Theor. Math. Phys. **19**, 923–954 (2015). arXiv:1208.6153 [math-ph]
325. M. Gunaydin, Exceptional realizations of lorentz group: supersymmetries and leptons. Nuovo Cim. **A29**, 467 (1975)
326. M. Gunaydin, Octonionic hilbert spaces, the poincare group and SU(3). J. Math. Phys. **17**, 1875 (1976)
327. F. Gursey, Octonionic structures in particle physics. Lect. Notes Phys. **94**, 508–521 (1979)
328. M. Gunaydin, Quadratic Jordan formulation of quantum mechanics and construction of lie (super)algebras from Jordan (super)algebras. In: International colloquium on group theoretical methods in physics, vol. 10, p. 18. Israel Grp.Th.Meth (1979)
329. G. Sierra, An application to the theories of Jordan algebras and freudenthal triple systems to particles and strings. Class. Quant. Gravit. **4**, 227 (1987)
330. F. Gursey, Superpoincare groups and division algebras. Mod. Phys. Lett. A **2**, 967 (1987)
331. M.J. Duff, Supermembranes: the first fifteen weeks. Class. Quant. Gravit. **5**, 189 (1988)
332. M. Blencowe, M.J. Duff, Supermembranes and the signature of space-time. Nucl. Phys. B **310**, 387 (1988)
333. M. Gunaydin, Generalized conformal and superconformal group actions and Jordan algebras. Mod. Phys. Lett. A **8**, 1407–1416 (1993). arXiv:hep-th/9301050 [hep-th]
334. N. Berkovits, A ten-dimensional super Yang–Mills action with off-shell supersymmetry. Phys. Lett. B **318**, 104–106 (1993). arXiv:hep-th/9308128 [hep-th]
335. C.A. Manogue, J. Schray, Finite Lorentz transformations, automorphisms, and division algebras. J. Math. Phys. **34**, 3746–3767 (1993). arXiv:hep-th/9302044

336. J.M. Evans, Auxiliary fields for super Yang–Mills from division algebras. Lect. Notes Phys. **447**, 218–223 (1995). arXiv:hep-th/9410239 [hep-th]
337. J. Schray, C.A. Manogue, Octonionic representations of Clifford algebras and triality. Found. Phys. **26**(1), 17–70 (1996). arXiv:hep-th/9407179
338. F. Gürsey, C.-H. Tze, On the role of division, Jordan and related algebras in particle physics. World Scientific, London (1996). http://www.worldscientific.com/doi/abs/10.1142/9789812819857
339. C.A. Manogue, T. Dray, Dimensional reduction. Mod. Phys. Lett. A **14**, 99–104 (1999). arXiv:hep-th/9807044 [hep-th]
340. E. Cremmer, B. Julia, H. Lu, C.N. Pope, Higher dimensional origin of D = 3 coset symmetries. arXiv:hep-th/9909099 [hep-th]
341. M. Günaydin, K. Koepsell, H. Nicolai, Conformal and quasiconformal realizations of exceptional Lie groups. Commun. Math. Phys. **221**, 57–76 (2001). arXiv:hep-th/0008063
342. F. Toppan, On the octonionic M-superalgebra. In: Sao Paulo 2002, Integrable theories, solitons and duality. (2002). arXiv:hep-th/0301163
343. M. Günaydin, O. Pavlyk, Generalized spacetimes defined by cubic forms and the minimal unitary realizations of their quasiconformal groups. JHEP **08**, 101 (2005). arXiv:hep-th/0506010
344. Z. Kuznetsova, F. Toppan, Superalgebras of (split-)division algebras and the split octonionic M-theory in (6,5)-signature. arXiv:hep-th/0610122
345. S. Bellucci, S. Ferrara, M. Günaydin, A. Marrani, Charge orbits of symmetric special geometries and attractors. Int. J. Mod. Phys. **A21**, 5043–5098 (2006). arXiv:hep-th/0606209
346. L. Borsten, D. Dahanayake, M.J. Duff, H. Ebrahim, W. Rubens, Black holes, qubits and octonions. Phys. Rep. **471**(3–4), 113–219 (2009). arXiv:0809.4685 [hep-th]
347. L. Borsten, D. Dahanayake, M.J. Duff, W. Rubens, Black holes admitting a Freudenthal dual. Phys. Rev. D **80**(2), 026003 (2009). arXiv:0903.5517 [hep-th]
348. J.C. Baez, J. Huerta, Division algebras and supersymmetry II. Adv. Theor. Math. Phys. **15**(5), 1373–1410 (2011). arXiv:1003.3436 [hep-th]
349. L. Borsten, D. Dahanayake, M.J. Duff, S. Ferrara, A. Marrani et al., Observations on integral and continuous u-duality orbits in N = 8 supergravity. Class. Quant. Gravit. **27**, 185003 (2010). arXiv:1002.4223 [hep-th]
350. M. Gunaydin, H. Samtleben, E. Sezgin, On the magical supergravities in six dimensions. Nucl. Phys. B **848**, 62–89 (2011). arXiv:1012.1818 [hep-th]
351. M. Rios, Extremal black holes as qudits. arXiv:1102.1193 [hep-th]
352. J. Huerta, Division algebras, supersymmetry and higher gauge theory. arXiv:1106.3385 [math-ph]
353. J. Huerta, Division algebras and supersymmetry III. Adv. Theor. Math. Phys. **16**, 1485–1589 (2012). arXiv:1109.3574 [hep-th]
354. S. Ferrara, A. Marrani, Black holes and groups of type E7. arXiv:1112.2664 [hep-th]
355. S.L. Cacciatori, B.L. Cerchiai, A. Marrani, Magic coset decompositions. Adv. Theor. Math. Phys. **17**, 1077–1128 (2013). arXiv:1201.6314 [hep-th]
356. A. Anastasiou, L. Borsten, M.J. Duff, L.J. Hughes, S. Nagy, An octonionic formulation of the M-theory algebra. JHEP **1411**, 022 (2014). arXiv:1402.4649 [hep-th]
357. J. Huerta, Division algebras and supersymmetry IV. arXiv:1409.4361 [hep-th]
358. A. Marrani, P. Truini, Exceptional lie algebras, SU(3) and Jordan pairs part 2: zorn-type representations. J. Phys. A **47**, 265202 (2014). arXiv:1403.5120 [math-ph]
359. A. Hurwitz, Uber die komposition der quadratishen formen von beliebig vielen variabeln. Nachr. Ges. Wiss. Gottingen 309–316 (1898)
360. R. Moufang, Alternativkörper und der satz vom vollständigen vierseit. Abh. Math. Sem. Hamb. **9**, 207–222 (1933)
361. P. Jordan, Über eine nicht-desarguessche ebene projektive geometrie. Abh. Math. Sem. Hamb. **16**, 74–76 (1949)
362. C. Chevalley, R.D. Schafer, The exceptional simple lie algebras f_4 and e_6. Proc. Natl. Acad. Sci. USA 137–141 (1950)
363. D.B. Fairlie, C.A. Manogue, A parametrization of the covariant superstring. Phys. Rev. D **36**, 475 (1987)
364. J. Schray, The general classical solution of the superparticle. Class. Quant. Gravit. **13**, 27–38 (1996). arXiv:hep-th/9407045 [hep-th]
365. T. Dray, J. Janesky, C.A. Manogue, Octonionic hermitian matrices with non-real eigenvalues. Adv. Appl. Clifford Algebras **10**(2), 193–216 (2000)

366. C. Hull, Symmetries and compactifications of (4,0) conformal gravity. JHEP **0012**, 007 (2000). arXiv:hep-th/0011215 [hep-th]
367. C. Hull, Conformal nongemetric gravity in six-dimensions and M theory above the Planck energy. Class. Quant. Gravit. **18**, 3233–3240 (2001). arXiv:hep-th/0011171 [hep-th]
368. Y.-T. Huang, A.E. Lipstein, Amplitudes of 3D and 6D maximal superconformal theories in supertwistor space. JHEP **1010**, 007 (2010). arXiv:1004.4735 [hep-th]
369. B. Czech, Y.-T. Huang, M. Rozali, Chiral three-point interactions in 5 and 6 dimensions. JHEP **1210**, 143 (2012). arXiv:1110.2791 [hep-th]