Queen Mary University of London

Thesis submitted for the degree of Doctor of Philosophy

New Structures in Gauge Theory and Gravity

Sam Wikeley

Supervisor

Dr Chris White

May 15, 2023

Centre for Theoretical Physics
School of Physical and Chemical Sciences
Queen Mary University of London

Declaration

I, Sam Wikeley, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: *Sam Wikeley*
Date: 15/05/2023

Details of collaboration and publications:

This thesis is based on the publications [1–4]. The work in [1–3] was carried out with my supervisor Chris White, along with Luigi Alfonsi in [1], Rashid Alawadhi and David Berman in [2], and Kymani Armstrong-Williams in [3]. The work in [4] was carried out with Ricardo Monteiro and Ricardo Stark-Muchão. Where other sources have been used, they are cited in the bibliography.

Abstract

Colour-kinematics duality provides new insights into the perturbative structure of quantum field theory. In particular, it recasts gravity as a double copy of gauge theory, an idea which has given rise to a variety of novel connections between these two seemingly disparate theories. In this thesis, we will explore a number of new examples of the double copy, which both extend the catalogue of cases in which it is known to apply and provide insights into theoretical structure of the correspondence. We will begin by investigating the role of non-local information in the double copy for classical solutions, leading to a topological condition that can be furnished with a double copy interpretation. As this condition is naturally expressed in terms of certain Wilson lines, we will go on to develop a double copy for the general form of these operators as well as the closely related geometrical concept of holonomy. We then further investigate the non-perturbative structure of the double copy by restricting to the self-dual sectors of gauge theory and gravity. Here we generalise the single copy structure of gravitational instantons, and provide new insights into the nature of the kinematic algebra underlying the double copy. Finally, we investigate the old idea that one-loop amplitudes in self-dual Yang-Mills and gravity are generated by an anomaly of the classical integrability of these theories. By writing explicit quantum-corrected actions for the self-dual theories, we will demonstrate a manifestation of this anomaly and uncover a novel double copy that holds at the level of the loop-integrated amplitudes.

Acknowledgements

I would like to thank my supervisor Chris White, who has influenced my thoughts about and attitude towards physics in an immeasurable way. I will always be grateful to Chris for his patience, encouragement, and seemingly endless enthusiasm. He has been a constant inspiration throughout my PhD and I am certain that it would not have been nearly as enjoyable without him as a supervisor.

My second supervisor, Ricardo Monteiro, also deserves great thanks. We embarked on a project that was stuck in the weeds for a long, long time, and without his perseverance I don't know if it would ever have seen the light of day. I am immensely grateful to him for his support and all that he has taught me.

I have been lucky to work with a number of brilliant physicists. I would like to thank my collaborators Rashid Alawadhi, Luigi Alfonsi, Kymani Armstrong-Williams, David Berman, and Ricardo Stark-Muchão. It has been a joy to work with you all.

The theory group at QMUL has been a wonderful environment within which to spend the past three-and-a-half years. I have, at this point, shared an office with many people, and I must thank them all for putting up with me. Thanks to George, Adrian, Bernie, David, Manuel, Stefano, Marcel, Shun-Qing, Linfeng, Rajath, Gergely, Enrico, Nadia, Graham, Josh, Lewis, Kymani, Chinmaya, Tancredi, Mitchell, Alex, Gus, George, Mahesh, and, of course, to Lara the dog, whose calming presence is unparalleled. Thank you also to Bill Spence and Anton Ilderton for conducting my viva and providing useful feedback on this thesis. Of these people, I must particularly thank George Barnes and Adrian Padellaro. We started the PhD together and I could not have asked for better people to go through this process with. It worries me to think of the amount of time I have spent distracting you both.

Finally, thank you to my family, my friends, and to Charlotte. Your love and support has made this possible.

This work was supported by the Science and Technology Facilities Council (STFC) Consolidated Grant ST/P000754/1 *String theory, gauge theory & duality.*

"Doubles the image, the two overlap, with the right sort of light, the right lenses, you can separate them in stages, a little further each time, step by step till in fact it becomes possible to saw somebody in half optically, and instead of two different pieces of one body, there are now two complete individuals walking around, who are identical in every way, capisci?"

Against the Day
THOMAS PYNCHON

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The past thirty years has seen an explosion of progress in the analytic computation of scattering amplitudes. On the one hand, there is experimental motivation. Precision computations in non-abelian gauge theories are critical for the success of large-scale scattering experiments such as the Large Hadron Collider (LHC), while the detection of gravitational waves by the LIGO and VIRGO collaborations in 2015 [6] has generated a demand for efficient approaches to perturbative gravity. Alongside the practical value that is inherent in the development of new tools for scattering amplitude calculations, there is great theoretical worth. In traditional Lagrangian-based approaches to scattering processes, the proliferation of Feynman diagrams as one goes to higher multiplicity and loop-order appears at odds with the relatively simple structure of the amplitudes themselves. This issue is particularly apparent in perturbative quantum gravity, for which one finds an infinite number of interaction vertices. By adopting a more on-shell philosophy, a hidden simplicity is revealed. Using the physical properties of the S-matrix such as locality, unitarity, and gauge invariance, a generic amplitude computation can be "bootstrapped" from its simplest constituent parts. Adopting this approach has given rise to an expansive toolkit of powerful on-shell techniques, such as BCFW recursion [7] and generalised unitarity [8], with which the computation of high multiplicity and loop-level amplitudes is vastly simplified.

As the modern amplitudes program has progressed, a number of deeply unexpected results have emerged. One of the most surprising is colour-kinematics duality, first discovered by Bern, Carrasco, and Johansson (BCJ) in refs. [9, 10]. Colour-kinematics duality brings into focus a highly non-trivial structure in gauge theory amplitudes that is almost completely obscured in traditional Lagrangian-based approaches. In refs. [9, 10], it was shown that in the computation of a non-abelian gauge theory scattering amplitude, the kinematic numerators that appear in each term can be arranged so as to obey identical algebraic relations to the colour factors. This is a surprising result. The colour factors are formed from contractions of structure constants of the gauge group Lie algebra, such that any identities they obey are inherited from the structure of the gauge group. The fact that the kinematics satisfy an analogous set of identities suggests the existence of a kinematic Lie algebra that is dual to the colour algebra, a structure that is obscured in the traditional approach to gauge theory. Furthermore, the existence of this duality suggests that sensible objects can be formed by replacing the colour factors in the gauge theory amplitude with a second set of kinematic factors. The resulting object, in which there are two copies of the kinematic factors, remarkably corresponds to a gravitational amplitude. This process, known as the BCJ double copy, offers a radically different perspective on the structure of gauge theory and gravity, recasting these two seemingly disparate theories in a strikingly similar language such that gluon amplitudes can be simply transformed into those of gravitons and related particles.

The first hints of a double copy-like structure can be traced back to the early days of string theory. In 1986, Kawaii, Lewellen, and Tye (KLT) observed that in bosonic string theory, tree-level closed string amplitudes can be expressed as sums over products of two tree-level open string amplitudes, where on either side of this equality the amplitudes considered are of the same multiplicity. In the low energy limit, the closed string amplitudes with massless spin-2 external states reduce to graviton amplitudes, while the open string amplitudes with spin-1 external states reduce to certain gluon amplitudes. These so-called KLT relations provide a correspondence between tree-level amplitudes in gauge theory and gravity at each multiplicity, which follows the rough schematic form of

$$\textit{gravity} = (\textit{gauge theory})^2. \tag{1.1}$$

While the KLT relations offer a suggestion of some deeper connection between gauge theory and gravity, they become increasingly cumbersome as one moves to higher multiplicity and do not permit a loop-level generalisation. Colour-kinematics duality and the BCJ double copy pick up the scent where the KLT relations trail off, providing a field theory realisation of the squaring relation of eq. (1.1) that is alluringly generalisable. At tree-level the BCJ double copy has been shown to reproduce the KLT relations [11], and has been proven to hold to all multiplicities via a variety of approaches [11–15]. However, the great power of the BCJ approach lies in its appearance at loop-level. While at present there is no general proof of its validity to all loop-orders, a wide variety of highly non-trivial examples have been demonstrated. We give a partial overview of these results in section 2.2.4.

The discovery of colour-kinematics duality and the bridge it provides between gauge theory and gravity raises many important questions. Is this merely a mathematical coincidence allowing us to cheat the computational formidability of gravity, or does it constitute some deeper connection between these two worlds that so persistently evade unification? If the answer is the former, the importance of the double copy would be in no way diminished. As a calculational tool, it has facilitated insights into sectors of gravity not previously accessible with traditional approaches, leading to a deeper understanding of the ultraviolet structure of gravity and supergravity theories (see refs. [5, 16] for extensive reviews). However, since its initial formulation at the level of gauge and gravity amplitudes, there has been a proliferation of examples of a double copy-like structure in domains and contexts vastly separated from one another. An expansive web of theories that are constructible from the principles of the double copy has emerged [5], and phenomena ranging from asymptotic symmetries [17] to solution-generating transformations [18, 19] to fluid dynamics [20] have been furnished with a double copy interpretation.

If the double copy is to be viewed as a statement about the fundamental nature of gauge and gravitational theories, it is natural to wonder whether its presence persists in some form at the level of classical physics. Indeed, it can be shown that perturbative classical solutions in gauge theory and gravity can be related by the double copy, as first demonstrated in ref. [21]. This is, in some sense, to be expected given the close relationship between perturbative solutions to equations of motion and tree-level amplitudes. The great surprise, however, is that for certain algebraically special solutions, the double copy can be applied at the level of *exact* classical solutions. The exact classical double copy was initially developed in ref. [21], where it was shown to hold for stationary vacuum spacetimes which permit a Kerr-Schild form. The basic example of this formalism explicitly links the Schwarzschild solution in pure general relativity with the Coulomb solution in linearised Yang-Mills theory via the replacement of kinematic with colour information, thereby establishing a double copy between the simplest point-particle configurations in gauge theory and gravity. Since this *Kerr-Schild double copy*, a number of generalisations have appeared. By working in the spinorial formalism, the *Weyl double copy* constitutes an exact equality between the Weyl spinor and two copies of the Maxwell spinor for all vacuum Type-D spacetimes [22], while further generalisations can be obtained from twistor space [23–25]. In section 2.3.4 we give a brief overview of the current status of the classical double copy.

There are many approaches to developing our understanding of colour-kinematics duality and the double copy. An obvious route is to extend the catalogue of examples in which the double copy can be shown to apply. At the level of amplitudes, this involves pushing the duality to higher-loop order and extending away from flat spacetime to curved backgrounds (see e.g. refs. [26–28]). For exact classical solutions, all current examples involve algebraically special solutions which linearise their equations of motion. By finding new examples of double copiable exact solutions, one might hope that a more general picture will emerge in which fully non-linear or strongly coupled solutions can be identified under the duality. In parallel to this line of work, it is important to examine the conceptual foundations of the double copy. A central question in this regard is the nature of the kinematic algebra. Explicit interpretations of the kinematic algebra are currently limited to special cases [29–34], the extension of which would be useful for both perturbative and non-perturbative studies. A more general understanding of the algebra would allow for efficient constructions of duality satisfying numerators for scattering amplitudes. Furthermore, at the level of exact classical solutions, the role of the kinematic algebra is more opaque. Bringing this structure into sharper focus might make clear when and why an exact double copy is possible.

In this thesis, based on refs. [1–4], we will make progress in these directions by developing novel manifestations of the double copy at both the classical and quantum level. Chapters 3, 4, and 5 will primarily focus on the non-perturbative structure of the correspondence. In chapter 3 we attempt to rectify the fact that all previous examples of the double copy involve local quantities. We use the known double copy relation between magnetic monopoles in gauge theory and the Taub-NUT solution in general relativity to investigate whether it is possible to identify *global* information on either side of the correspondence. We find that when phrased in terms of a certain patching condition, the topological classification of these solutions satisfies the same double copy structure as the solutions themselves, such that the map takes the form of a local and global statement. The patching conditions will be written in terms of Wilson line operators, whose general double copy interpretation we develop in chapter 4. In doing so, the question of how to treat the geometrical concept of holonomy from the perspective of the double copy will naturally arise. We will find a single copy of the gravitational holonomy, thus completing a square of four operators that can be related by the double copy.

The double copy is particularly well understood in the self-dual sectors of gauge theory and gravity. We use this fact to gain insights into both non-perturbative and perturbative aspects of the duality in chapters 5 and 6 respectively. In chapter 5 we extend the known double copy structure of gravitational instantons, through which we will gain insights into the role of the kinematic algebra at the level of exact classical solutions. In chapter 6, we investigate an old proposal that the one-loop amplitudes in self-dual Yang-Mills and gravity are generated by the anomaly of the classical integrability of these theories. By developing a formalism that manifests this idea we will uncover a novel double copy between loop-integrated one-loop effective vertices. This provides the first example of a loop-level double copy that holds at the level of the loop-integrated amplitude, rather than the loop-integrand.

To begin however, we provide in the following chapter a brief overview of the basics of colour-kinematics duality and the double copy, along with a discussion of the current status of the field.

Chapter 2

Colour-kinematics duality and the double copy

2.1 Scattering amplitudes in gauge theory and gravity

In this section we will review the basics of scattering in gauge theory and gravity. To begin, we will briefly discuss the diagrammatic approach to the computation of amplitudes. This will allow us to identify the mathematical structures of interest in the BCJ double copy for amplitudes, and to fix notation and conventions that will be used throughout this work. We use the $(-+++)$ metric signature throughout, except where stated otherwise.

2.1.1 Gauge theory

Pure Yang-Mills theory is an interacting theory of massless spin-1 fields with a local non-abelian symmetry group. In d-dimensions, it is described by the action

$$S_{\mathrm{YM}} = -\frac{1}{4}\int d^d x\, \mathrm{tr}(F^{\mu\nu}F_{\mu\nu}), \tag{2.1}$$

where $F_{\mu\nu}$ is the field strength, defined in terms of the gauge field A_μ via

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu], \tag{2.2}$$

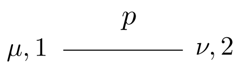
and the trace is over the colour degrees of freedom. The gauge field is valued in the Lie algebra of the gauge group G, such that it may be expanded as $A_\mu = A^a_\mu T^a$ where T^a

are the generators of G. Note that eq. (2.2) implies that the field strength is also Lie algebra valued, $F_{\mu\nu} = F^a_{\mu\nu} T^a$. The generators are traceless and Hermitian, and satisfy the following relations:

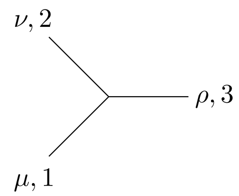
$$[T^a, T^b] = i f^{abc} T^c, \qquad \mathrm{tr}(T^a T^b) = C_r \delta^{ab}. \tag{2.3}$$

Here the first equation defines the algebra with structure constants f^{abc}, while the second is a normalisation condition with C_r a representation-dependent constant. For now we will consider the gauge group to be $G = \mathrm{SU}(N)$, for which there are $N^2 - 1$ generators.

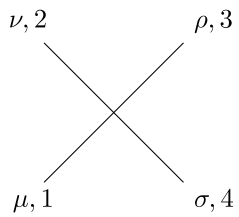
To compute scattering amplitudes in the textbook action-based approach (see e.g. ref. [35]), we need to fix a gauge and extract a set Feynman rules from the action in eq. (2.1). In the Feynman gauge we find the following set of rules [35]:



$$= -\frac{i\eta_{\mu\nu}\delta^{a_1a_2}}{p^2} \tag{2.4}$$



$$= g f^{a_1a_2a_3}\left[\eta_{\mu\nu}(p_1 - p_2)_\rho + \text{cyclic}\right] \tag{2.5}$$



$$\begin{aligned} = -ig^2 \Big[& f^{a_1a_2b} f^{ba_3a_4}(\eta_{\mu\rho}\eta_{\nu\sigma} - \eta_{\mu\sigma}\eta_{\nu\rho}) \\ & + f^{a_1a_3b} f^{ba_2a_4}(\eta_{\mu\nu}\eta_{\rho\sigma} - \eta_{\mu\sigma}\eta_{\nu\rho}) \\ & + f^{a_1a_4b} f^{ba_2a_3}(\eta_{\mu\nu}\eta_{\rho\sigma} - \eta_{\mu\rho}\eta_{\nu\sigma}) \Big] \end{aligned} \tag{2.6}$$

Here we have a propagator, three-point vertex, and four-point vertex respectively. With these rules integrands for each diagrammatic contribution can be constructed, following which any loop-momenta must be integrated over. The important point, however, is that in principle it is only this set of three rules that is required to compute Yang-Mills amplitudes to all multiplicities and at arbitrary loop-level.

A generic Feynman diagram formed from the above rules will always feature a colour factor, arising from contractions of the gauge group structure constants that appear in the vertices. By finding a basis that spans the colour degrees of freedom of an amplitude, we can disentangle the colour and kinematics in a way that will greatly simplify calculations. To start, let us rewrite the structure constants in terms of the gauge group generators by using the relation

$$f^{abc} = -\frac{i}{C_r}\,\mathrm{tr}(T^a[T^b, T^c]). \tag{2.7}$$

A generic string of structure constants will now be a product of traces of generators. Contractions over adjoint indices can then be performed using the completeness relation for the SU(N) generators:

$$(T^a)^i{}_j (T^a)^k{}_l = \delta^i_l \delta^k_j - \frac{1}{N}\delta^i_j \delta^k_l. \tag{2.8}$$

Performing all such contractions will reduce the colour structure of any tree-level amplitude to be only in terms of single traces of generators. It is then possible to write the n-point tree-level amplitudes in pure Yang-Mills as

$$\mathcal{A}_n^{\mathrm{tree}} = g^{n-2} \sum_{\text{perms }\sigma} \mathrm{tr}(T^{a_1} T^{a_{\sigma(2)}} \cdots T^{a_{\sigma(n)}}) A_n[1, \sigma(2), ..., \sigma(n)]. \tag{2.9}$$

We have therefore found a basis that spans the colour degrees of freedom of the amplitude, as desired. At loop-level a similar decomposition is possible but it will also contain multi-traces (see e.g. ref. [36]). The sum is over the basis of $(n-1)!$ elements that takes into account the cyclic property of the traces. The coefficients A_n of the colour basis are *colour-ordered* or *partial amplitudes*. They are gauge invariant objects that are dependent only on kinematic information. Furthermore, their computation is simpler than the full amplitudes as they only receive contributions from fixed cyclic orderings of external legs.

2.1.2 Perturbative gravity

Consider the d-dimensional Einstein-Hilbert action with zero cosmological constant

$$S_{\mathrm{EH}} = \frac{2}{\kappa^2}\int d^dx \sqrt{-g}\, R. \tag{2.10}$$

Here κ is the gravitational coupling, related to Newton's constant via $\kappa = \sqrt{32\pi G_N}$, g is the determinant of the metric, and R is the Ricci scalar. This action describes pure general relativity, and thus yields the Einstein field equations after variation with respect to the metric.

We wish to study flat-space scattering amplitudes in an analogous manner to the Yang-Mills approach; that is, by extracting a set of Feynman rules from the Einstein-Hilbert action. This can be done in the weak-field limit by expanding the metric as a perturbation around flat spacetime:

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}. \tag{2.11}$$

The fluctuation $h_{\mu\nu}$ is referred to as the graviton field. Inserting this expansion into the action in eq. (2.10) we obtain a theory describing the interaction of perturbative graviton states at weak coupling. The action can be written schematically as [16]

$$S_{\text{EH}} \sim \int d^d x \left(h\partial^2 h + \kappa h^2 \partial h + \kappa^2 h^3 \partial h + ... \right). \tag{2.12}$$

This theory poses a number of problems for the diagrammatic approach to the computation of scattering amplitudes. Firstly, the expanded action contains an infinite number of interaction vertices of arbitrary multiplicity, as denoted by the ellipsis in eq. (2.12). This is due to the expansions of both the inverse metric in the Ricci scalar and the square root of the metric determinant. Note that this situation is in stark contrast to the two vertices with a maximum of four external legs that appeared in Yang-Mills theory. Furthermore, the structure of each individual vertex is exceedingly complex, even for low numbers of external legs. This can be seen already in the three-point vertex. Adopting the de Donder gauge, the three-point graviton vertex is [37]

$$\begin{aligned} G^{(3)}_{\mu\rho,\nu\lambda,\sigma\tau}(p_1,p_2,p_3) \sim i\kappa \,\text{Sym}\Big[& -\frac{1}{2} P_3 \left(p_1 \cdot p_2 \eta_{\mu\rho} \eta_{\nu\lambda} \eta_{\sigma\tau} \right) - \frac{1}{2} P_6 \left(p_{1\nu} p_{1\lambda} \eta_{\mu\rho} \eta_{\sigma\tau} \right) \\ & + \frac{1}{2} P_3 \left(p_1 \cdot p_2 \eta_{\mu\nu} \eta_{\rho\lambda} \eta_{\sigma\tau} \right) + P_6 \left(p_1 \cdot p_2 \eta_{\mu\rho} \eta_{\nu\sigma} \eta_{\lambda\tau} \right) \\ & + 2P_3 \left(p_{1\nu} p_{1\tau} \eta_{\mu\rho} \eta_{\lambda\sigma} \right) - P_3 \left(p_{1\lambda} p_{2\mu} \eta_{\rho\nu} \eta_{\sigma\tau} \right) \\ & + P_3 \left(p_{1\sigma} p_{2\tau} \eta_{\mu\nu} \eta_{\rho\lambda} \right) + P_6 \left(p_{1\sigma} p_{1\tau} \eta_{\mu\nu} \eta_{\rho\lambda} \right) \\ & + 2P_6 \left(p_{1\nu} p_{2\tau} \eta_{\lambda\mu} \eta_{\rho\sigma} \right) + 2P_3 \left(p_{1\nu} p_{2\mu} \eta_{\lambda\sigma} \eta_{\tau\rho} \right) \\ & - 2P_3 \left(p_1 \cdot p_2 \eta_{\rho\nu} \eta_{\lambda\sigma} \eta_{\tau\mu} \right) \Big]. \end{aligned} \tag{2.13}$$

Here Sym denotes a symmetrisation in each pair of graviton Lorentz indices $(\mu\rho)$, $(\nu\lambda)$, and $(\sigma\tau)$, and P_i denotes a summation over distinct permutations of the external momenta, with the subscript i indicating the number of terms generated in each case. The complexity of this vertex as compared to the three-point Yang-Mills vertex in eq. (2.5) is clear. Combining this with the fact that the theory contains an infinite number of additional vertices, each with increasing complexity, makes the computation of gravitational scattering amplitudes using Feynman diagrams incredibly challenging, even at low multiplicity. As the number of external legs increases and loops are considered, computations based on the diagrammatic expansion become unfeasible.

2.1.3 Amplitudes without actions

In the modern approach to scattering amplitudes one rarely takes the approach outlined above. A broad set of techniques have been developed to circumvent the use of Feynman diagrams or even the need for an action at all for the computation of amplitudes (see e.g. refs. [16, 38] for reviews). The challenges encountered above can be traced back to the fact that there is a enormous amount of redundancy in the definition of an action. Actions are written in terms of off-shell fields, often with gauge or diffeomorphism symmetries, that are indifferent to arbitrary field redefinitions. However, the end-goal is to compute gauge-invariant on-shell scattering amplitudes, whose form is usually far simpler than that of the Feynman diagrams that enter their computation. The philosophy of the modern amplitudes program is to circumvent these redundancies by constraining the form of amplitudes by the physical properties of the S-matrix itself. Reformulating amplitude calculations in this way results in a profoundly different approach to perturbative quantum field theory as a whole, while also revealing the presence of novel mathematical structures in the amplitudes themselves.

We can immediately see a manifestation of this principle in the three-point graviton vertex in eq. (2.13). This is a gauge-dependent object written in terms of off-shell momenta. Let us now remove this gauge dependency by contracting the vertex into physical on-shell states satisfying

$$p_{i\,\mu}\epsilon_i^{\mu\nu} = 0, \quad \epsilon_i^{\mu\nu} = \epsilon_i^{\nu\mu}, \quad \eta_{\mu\nu}\epsilon^{\mu\nu} = 0, \tag{2.14}$$

where $\epsilon_i^{\mu\nu}$ is the polarisation tensor associated with leg i. Equation (2.13) then reduces to

$$G^{(3)}_{\mu\rho,\nu\lambda,\sigma\tau}(p_1,p_2,p_3) \sim -i\kappa\big[(p_1-p_2)_\sigma\eta_{\mu\nu} + \text{cyclic}\big]\big[(p_1-p_2)_\tau\eta_{\rho\lambda} + \text{cyclic}\big]. \tag{2.15}$$

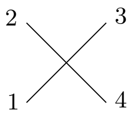
This expression is clearly far more manageable than the off-shell vertex in eq. (2.13). However, more interestingly, in taking the vertex on-shell we have unearthed a relation to the three-point gauge theory vertex in eq. (2.5). Ignoring constant factors, the on-shell graviton vertex looks like the three-point gauge theory vertex but with the colour structure replaced by a second copy of the kinematic factor. This similarity between the vertices was completely obscured in the off-shell formulation.

2.2 Colour-kinematics duality and the BCJ double copy

2.2.1 A first look at colour-kinematics duality

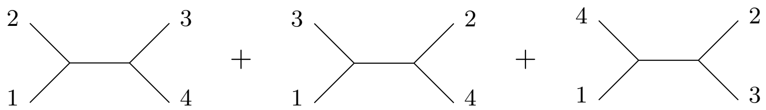
In the previous section we saw that when written on-shell the three-point graviton vertex is related to the three-point gluon vertex via some replacement of colour information with kinematic information. This is the driving idea behind the double copy, and it is made possible due to a principle known as colour-kinematics duality. Before discussing these concepts in generality, it will be useful to see them at play in a simple, but physical, example: the four-point tree-level amplitude.

The four-point tree-level amplitude in Yang-Mills theory can be computed with the Feynman rules in eqs. (2.4) - (2.6) (see e.g. ref. [35]). The calculation involves four diagrams: three that are constructed from the cubic vertex and the propagator, giving rise to the s-, t-, and u-channel contributions in figure 2.1, and one from the on-shell four-point vertex. The four-point vertex can itself be written in terms of s-, t-, and u-channel cubic contributions via trivial multiplications of $1 = s/s = t/t = u/u$. This can be seen as follows



$$\sim f^{a_1 a_2 b} f^{b a_3 a_4} n_s^{(4)} + f^{a_1 a_3 b} f^{b a_2 a_4} n_u^{(4)} + f^{a_1 a_4 b} f^{b a_2 a_3} n_t^{(4)}$$

$$= f^{a_1 a_2 b} f^{b a_3 a_4} n_s^{(4)} \left(\frac{s}{s}\right) + f^{a_1 a_3 b} f^{b a_2 a_4} n_u^{(4)} \left(\frac{u}{u}\right) + f^{a_1 a_4 b} f^{b a_2 a_3} n_t^{(4)} \left(\frac{t}{t}\right)$$

$\sim$ 

, (2.16)

where $n_i^{(4)}$ represents the kinematic coefficient of each colour factor in the vertex. Adding this to the contributions from the diagrams in figure 2.1, and grouping together terms with same colour factor yields the amplitude in the form

$$i\mathcal{A}_4 = g^2 \left(\frac{c_s n_s}{s} + \frac{c_t n_t}{t} + \frac{c_u n_u}{u}\right). \tag{2.17}$$

Here c_i and n_i refer to the colour and kinematic factors respectively. The s, t, and u subscripts refer to the channels shown in figure 2.1, and in the denominators are Mandelstam variables corresponding to each channel:

$$s = (p_1 + p_2)^2, \quad t = (p_1 + p_4)^2, \quad u = (p_1 + p_3)^2. \tag{2.18}$$

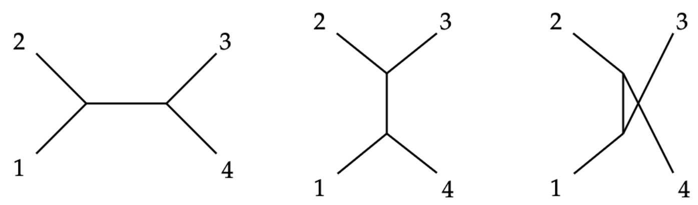


Figure 2.1: The s-, t-, and u-channel contributions respectively.

The three colour factors c_i in eq. (2.17) are given by

$$c_s = f^{a_1 a_2 b} f^{b a_3 a_4}, \quad c_t = f^{a_1 a_4 b} f^{b a_2 a_3}, \quad c_u = f^{a_1 a_3 b} f^{b a_2 a_4}. \tag{2.19}$$

Note that these colour factors sum to zero as a consequence of the the Jacobi identity of the gauge group algebra

$$c_s + c_t + c_u = 0. \tag{2.20}$$

The three kinematic factors n_i contain the kinematic information corresponding to each channel, apart from that associated with the propagators. They are formed from various contractions of momenta and polarisation vectors, satisfying the on-shell conditions $p_i^2 = 0$, $p_i \cdot \epsilon_i = 0$. The n_s factor is given by

$$\begin{aligned} n_s = &\left[(\epsilon_1 \cdot \epsilon_2)\, p_1^\mu + 2\,(\epsilon_1 \cdot p_2)\, \epsilon_2^\mu - (1 \leftrightarrow 2)\right] \left[(\epsilon_3 \cdot \epsilon_4)\, p_{3\mu} + 2\,(\epsilon_3 \cdot p_4)\, \epsilon_{4\mu} - (3 \leftrightarrow 4)\right] \\ &+ s\left[(\epsilon_1 \cdot \epsilon_3)\,(\epsilon_2 \cdot \epsilon_4) - (\epsilon_1 \cdot \epsilon_4)\,(\epsilon_2 \cdot \epsilon_3)\right], \end{aligned} \tag{2.21}$$

while the n_t and n_u factors are obtained via the replacements

$$n_t = n_s\big|_{1\to2\to3\to1}, \quad n_u = n_s\big|_{1\to3\to2\to1}. \tag{2.22}$$

Surprisingly, the sum of the three kinematic factors is zero,

$$n_s + n_t + n_u = 0, \tag{2.23}$$

in direct analogy with the colour factors in eq. (2.20). This relation is an example of a *kinematic Jacobi identity*. This is an unexpected result. The constraint on the colour factors in eq. (2.20) is obtained from purely group theoretic arguments originating from the Lie algebra structure of the gauge group. At face value, there is no reason to expect the kinematic dependence of the amplitude to satisfy an analogous constraint. One might think that this is nothing but a peculiarity of this simple example. However, the presence of this correspondence between the colour and kinematic degrees of freedom in the four-point tree-level Yang-Mills amplitude is a primitive example of a more general statement known as colour-kinematics duality, which we will discuss in more detail in the next section.

The four-point Yang-Mills amplitude should be a gauge invariant quantity, such that it is invariant under the transformation $\epsilon_i \to \epsilon_i + p_i$. To check that this is the case, let us confirm that eq. (2.17) vanishes under $\epsilon_i \to p_i$. Consider applying this transformation to leg four in eq. (2.21), such that

$$\begin{aligned} n_s &\to s\left[(\epsilon_1 \cdot \epsilon_2)(\epsilon_3 \cdot p_2 - \epsilon_3 \cdot p_1) + \text{cyclic}(1,2,3)\right] \\ &= s\alpha_s. \end{aligned} \tag{2.24}$$

The function α_s is invariant under cyclic permutations of $(1, 2, 3)$. As the other kinematic numerators are obtained via the cyclic replacements in eq. (2.22), the analogous functions obtained in the transformation of these numerators will all be equal, such that $\alpha_s = \alpha_t = \alpha_u \equiv \alpha$. Thus, under the transformation $\epsilon_4 \to p_4$, the amplitude reduces to

$$\begin{aligned} i\mathcal{A}_4 &\to g^2\left(c_s\alpha_s + c_t\alpha_t + c_u\alpha_u\right) \\ &= (c_s + c_t + c_u)\alpha, \end{aligned} \tag{2.25}$$

This vanishes as a consequence of the Jacobi identity in eq. (2.20), such that the amplitude is gauge invariant.

The fact that the colour and kinematic factors in the four-point amplitude satisfy identical algebraic relations, eqs. (2.20) and (2.23), suggests that they are interchangeable. That is, by replacing a colour factor with a kinematic factor or vice versa should result in a well-defined object. Consider the following replacements in eq. (2.17):

$$i\mathcal{A}_4\Big|_{\substack{c_i \to n_i \\ g \to \kappa/2}} \equiv i\mathcal{M}_4 = \left(\frac{\kappa}{2}\right)^2 \left(\frac{n_s^2}{s} + \frac{n_t^2}{t} + \frac{n_u^2}{u}\right). \tag{2.26}$$

The new object $\mathcal{M}_4$ is referred to as the *double copy* of $\mathcal{A}_4$, as the replacement $c_i \to n_i$ has resulted in a doubling of each kinematic numerator. What then does $\mathcal{M}_4$ correspond to? Remarkably, the answer turns out to be a four-point tree-level scattering amplitude in a gravitational theory [39]. We will not formally prove this fact here, but instead discuss why this should intuitively be the case.

Firstly, if eq. (2.26) is to be a gravitational amplitude it should arise from the scattering of external graviton states. The squaring of the kinematic numerators implies that the polarisation vectors can be promoted to symmetric polarisation tensors $\epsilon^{\mu\nu} = \epsilon^\mu \epsilon^\nu$. For null polarisation vectors we obtain a traceless $\epsilon^{\mu\nu}$, and thus we find the traceless symmetric polarisation tensors associated with graviton external states. Furthermore, we require that the amplitude obtained via the double copy is gauge invariant. By gauge invariance in perturbative gravity we mean invariance under linearised diffeomorphisms of the graviton field. Such a transformation takes the form

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu, \tag{2.27}$$

where the function ξ_μ parameterises the diffeomorphism. This implies that the amplitude should vanish under the following transformations of the polarisation tensors

$$\epsilon_i^{\mu\nu} \to p_i^\mu \epsilon_i^\nu + p_i^\nu \epsilon_i^\mu. \tag{2.28}$$

Applying this transformation to leg 4 we find

$$i\mathcal{M}_4 \sim (n_s + n_t + n_u)\, \alpha \tag{2.29}$$

which vanishes as a consequence of eq. (2.23). Thus, it is the fact that the kinematic numerators satisfy a kinematic Jacobi identity that enforces the invariance of the amplitude under linearised diffeomorphisms, in exact analogy to the colour factors in gauge theory.

2.2.2 Colour-kinematics duality

In the previous section we reviewed a simple manifestation of colour-kinematics duality and discussed how its presence allowed us to obtain the tree-level four-point amplitude in perturbative gravity by "double copying" the analogous Yang-Mills amplitude. This is the simplest physical example of a far more general story, which we turn to now.

For our purposes, it will be sufficient to consider pure gauge theory, containing only massless adjoint-valued fields. Consider then a general L-loop, m-point scattering amplitude in d-dimensional pure gauge theory:

$$\mathcal{A}_m^{(L)} = i^{L-1} g^{m-2+2L} \sum_{i\in\Gamma} \int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{1}{S_i} \frac{c_i n_i}{D_i}. \tag{2.30}$$

Here g is the coupling and the sum is over all distinct cubic graphs Γ. Associated with each graph is a symmetry factor S_i, a set of loop momenta k_l, an inverse factor of D_i containing a product of the internal Feynman propagator denominators, a colour factor c_i, and a kinematic factor n_i. The integral is over all loop-momenta. As we have seen in the example in the previous section, the colour and kinematic factors are the main players in colour-kinematics duality, so let us discuss them in more detail.

The colour factors c_i contain all information pertaining to the non-abelian gauge group G under which the fields in the theory transform. As we are considering pure gauge theory, a generic colour factor will be formed from a string of contracted Lie algebra structure constants f^{abc}. Thus, while each diagram in the sum in eq. (2.30) corresponds to a unique colour factor, these factors will not in general be independent, and will satisfy linear relations inherited from the structure constants, such as the Jacobi identity

$$f^{abe}f^{ecd} + f^{bce}f^{ead} + f^{cae}f^{ebd} = 0. \tag{2.31}$$

There will therefore exist triplets of graphs $\{i, j, k\}$ which satisfy the identity

$$c_i + c_j + c_k = 0, \tag{2.32}$$

which we refer to as a *colour Jacobi identity.*

The kinematic degrees of freedom of the amplitude are contained within the kinematic numerators n_i and the propagator factors D_i. A generic kinematic numerator will be a gauge-dependent function of the kinematic data associated with a given graph, such as the external momenta, loop momenta, and polarisations. A set of graphs $\{i, j, k\}$ are said to obey *colour-kinematics duality* or *BCJ duality* if their kinematic numerators have the same algebraic properties as their colour numerators. That is, for a set of duality satisfying diagrams $\{i, j, k\}$ we have

$$c_i + c_j + c_k = 0 \iff n_i + n_j + n_k = 0, \tag{2.33}$$

$$c_i = -c_j \iff n_i = -n_j. \tag{2.34}$$

In eq. (2.33), the identity satisfied by the kinematic numerators is referred to as a *kinematic Jacobi identity.* The second relation, in eq. (2.34), states that when a colour factor is antisymmetric, for example under the interchange of two legs, the corresponding kinematic factor is similarly antisymmetric. These are the defining relations of colour-kinematics duality.

The existence of colour-kinematics duality in an amplitude is far from obvious as kinematic numerators are non-unique and will in general not satisfy the relations of eqs. (2.33) and (2.34). As stated already, the kinematic numerators are gauge-dependent objects, however it turns out they enjoy a far greater freedom than this alone. Note that the amplitude in eq. (2.30) is invariant under the transformations

$$n_i \to n_i' = n_i + \Delta_i, \tag{2.35}$$

where $\{\Delta_i\}$ are arbitrary functions of kinematic variables subject to the constraint

$$\sum_{i\in\Gamma}\int\prod_{l=1}^{L}\frac{d^dk_l}{(2\pi)^d}\frac{1}{S_i}\frac{\Delta_i c_i}{D_i}=0. \tag{2.36}$$

This is referred to as a *generalised gauge transformation.* Thus, even if the numerators $\{n_i\}$ satisfy the duality relations in eqs. (2.33) and (2.34), there is an enormous amount of choice in writing the amplitude in terms of $\{n_i'\}$ for which colour-kinematics duality will no longer be apparent.

Colour-kinematics duality states that for a generic amplitude in pure non-abelian gauge theory it is always possible to find a representation in which its colour and kinematic numerators satisfy the same algebraic relations, as in eqs. (2.33) and (2.34). This has been proven at tree-level to all m [11], however it remains a conjecture at loop-level. Despite this, a large number of loop-level examples of duality satisfying amplitudes have been found, a selection of which we review in section 2.2.4.

2.2.3 The BCJ double copy

A remarkable consequence of colour-kinematics duality is that of the *BCJ double copy.* Consider a gauge theory amplitude as in eq. (2.30). The double copy construction states that given a set of kinematic numerators $\{\tilde{n}_i\}$ which satisfy colour-kinematics duality, a gravitational scattering amplitude can be obtained from eq. (2.30) via the simple replacements

$$c_i\to\tilde{n}_i,\quad g\to\frac{\kappa}{2}, \tag{2.37}$$

These are valid replacements as the numerators $\{\tilde{n}_i\}$ obey identical algebraic properties to the colour factors as a result of the duality, and the induced change in dimensionality is compensated for by the replacement of the coupling. Performing these replacements in eq. (2.30) yields

$$\mathcal{M}_m^{(L)}=i^{L-1}\left(\frac{\kappa}{2}\right)^{m-2+2L}\sum_{i\in\Gamma}\int\prod_{l=1}^{L}\frac{d^dk_l}{(2\pi)^d}\frac{1}{S_i}\frac{n_i\tilde{n}_i}{D_i}. \tag{2.38}$$

This object describes an m-point, L-loop scattering amplitude in a gravitational theory [9,10], and is referred to as the double copy of the gauge theory amplitude in (2.30). Furthermore, starting from the gravitational amplitude it is possible to obtain the gauge theory amplitude by performing the inverse replacements to those in (2.37). This process is referred to as taking the *single copy.*

There is an important subtlety in making the double copy replacements of eq. (2.37). Depending on the gauge group under consideration, it may be that certain contractions of structure constants produce new algebraic relations between colour factors that are specific to that gauge group. Such contractions can be thought of as analogous to performing the loop integrals in the kinematic factors. As the double copy is taking place in the integrand it is important that the colour factors are kept general, such that colour and kinematics are treated on an equal footing.

In the construction of the gravitational amplitude in eq. (2.38), very little was specified about the properties of the two sets of kinematic numerators $\{n_i\}$ and $\{\tilde{n}_i\}$. All that was required is that they describe two valid representations of the gauge theory amplitude, and that one set satisfies colour-kinematics duality. This freedom has important consequences for the structure of the gravitational theories obtained by the double copy. Firstly, the two sets of kinematic numerators are allowed to describe different external states. For example, the spectrum of states in the gravitational theory that arises when both sets of kinematic numerators are taken from 4-dimensional pure Yang-Mills will be obtained from tensor products of gluon states. In this case we will find the graviton

$$\text{graviton}^{\pm 2} = \text{gluon}^{\pm 1} \otimes \text{gluon}^{\pm 1}, \tag{2.39}$$

but also two scalar states from combinations of opposite helicity gluons, corresponding to a dilaton and an axion for the symmetric and antisymmetric combinations respectively:

$$\left.\begin{matrix}\text{dilaton}\\ \text{axion}\end{matrix}\right\} = \text{gluon}^{\pm 1} \otimes \text{gluon}^{\mp 1} \tag{2.40}$$

The double copy of pure Yang-Mills theory with itself is therefore not pure general relativity, but a gravitational theory containing a graviton, dilaton, and axion. Such a theory appears naturally in the low-energy limit of bosonic string theory, and is referred to as $\mathcal{N} = 0$ supergravity or NS-NS gravity.

Furthermore, there is no requirement that the two sets of kinematic numerators in eq. (2.38) come from amplitudes in the same gauge theory. This freedom allows for a wide variety of gravitational theories to be obtained by taking different gauge theories as input into the double copy. This has led to a vast web of double copy constructible theories, as reviewed in ref. [5].

As a small example of this web of theories, let us consider one example of another theory that is related to gauge theory and gravity due to colour-kinematics duality, and that will play a role later in this thesis. For a set of duality satisfying numerators it should be possible to obtain a sensible gauge invariant object by moving in the opposite direction

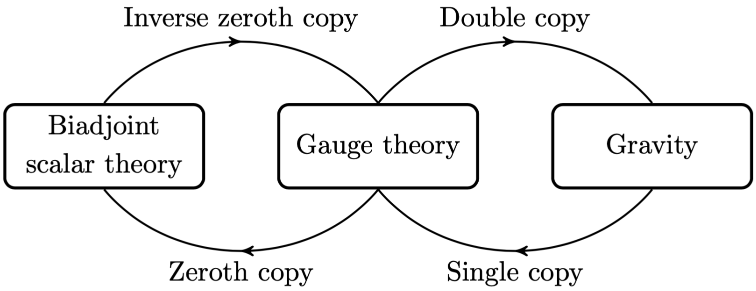


Figure 2.2: Schematic depiction of three theories related by colour-kinematics duality. This ladder of theories is itself a subset of a far greater web of double copy constructible theories, as reviewed in ref. [5].

to the double copy and replacing the kinematic numerator with a second colour factor. To this end, consider making the following replacements in the gauge theory amplitude in eq. (2.30):

$$n_i \to \tilde{c}_i, \quad g \to y, \tag{2.41}$$

where y is a new coupling parameter. These replacements are often referred to as taking the *zeroth copy*, and the resulting object is

$$\mathcal{T}_m^{(L)} = i^{L-1} y^{m-2+2L} \sum_{i\in\Gamma} \int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{1}{S_i} \frac{c_i \tilde{c}_i}{D_i}. \tag{2.42}$$

This describes a general scattering amplitude in a *biadjoint scalar theory*, containing scalar fields $\Phi^{aa'}$ which transform in the adjoint representation of two, possibly different, Lie algebras [21, 40, 41]. We are thus left with a chain of maps between three seemingly disparate theories made possible by the presence of colour-kinematics duality and the double copy construction, as is shown schematically in figure 2.2.

2.2.4 The BCJ double copy today

The BCJ double copy for amplitudes was first developed in refs. [9, 10]. The existence of numerators satisfying colour-kinematics duality in pure Yang-Mills has since been proven at tree-level to all multiplicity via a variety of approaches [11–15]. Furthermore, for such numerators a set of relations between n-point colour-ordered amplitudes, known as BCJ relations, can be derived [9]. These BCJ relations can themselves be independently derived from both string theory [42–44] and field theory [45]. One can then reverse the logic and show that the algebraic relations amongst kinematic numerators follow from the existence of the BCJ relations and Kleiss-Kuijf identities [40]. With the existence of duality satisfying numerators at tree-level on firm ground, the question is then whether the application of the double copy replacement rules always yields the desired gravitational amplitude. This has been proven to be the case provided one starts with a set of local numerators in Yang-Mills theory, via an inductive construction comparing the gravity amplitudes obtained via the double copy and BCFW recursion [39].

At loop-level there is currently no general proof that duality satisfying numerators always exist in Yang-Mills. Despite this, a vast array of highly non-trivial examples of the double copy in action at loop-level have been found, a few of which we highlight here. Perhaps the most general loop-level constructions manifesting colour-kinematics duality are the one-loop n-point amplitudes in pure Yang-Mills with all-plus and single-minus external particle helicity configurations [46]. We will return to these amplitudes in chapter 6. By introducing supersymmetry, we can increase the loop-order. In $\mathcal{N} = 4$ super-Yang-Mills duality satisfying numerators have been constructed up to four loops at $n = 4$ [47], up to two loops at $n = 5$ [48], and at one loop up to $n = 7$ [49]. Many other loop-level examples exist as well as extensions to more exotic theories, and we refer to ref. [5] for a comprehensive review of these developments.

The traditional approach to scattering amplitudes involves writing down a Lagrangian and extracting Feynman rules. One might wonder then whether it is possible to construct Lagrangians whose Feynman rules automatically yield duality-satisfying numerators. This poses a challenge as Lagrangians are inherently off-shell and so we lose much of the on-shell simplicity of the modern amplitudes programme. Despite this much progress has been made. Lagrangians manifesting colour-kinematics duality have been found up to six-points [39, 50] and to any multiplicity in the next-to-MHV sector [51]. Cubic-order Lagrangians of this kind were constructed in refs. [52, 53]. A Lagrangian-level double copy analysis of the BRST formalism has been developed in refs. [54, 55].

2.3 The double copy for classical solutions

In the preceding section the BCJ double copy for scattering amplitudes was introduced. This is an intrinsically perturbative construction, that relies on colour-kinematics duality holding for the colour and kinematic factors that appear in the perturbative expansion of an amplitude. However, given the mounting evidence for the validity of colour-kinematics duality at loop-level, it is natural to ask whether the double copy also holds at the classical level. That this should be possible is not immediately obvious as scattering amplitudes are by definition gauge independent objects while general classical solutions are not. The invariance of an amplitude under generalised gauge invariance plays an important role in finding duality satisfying representations of kinematic numerators, and thus identifying when a double copy is possible. Furthermore, it is unclear what the replacement rules should be in a classical solution as we no longer have a neat general decomposition into colour numerators, kinematic numerators, and propagators, as for amplitudes. Despite this, a broad variety of classical solutions, both perturbative and exact, in gauge theory and gravity have been related under the double copy philosophy of replacing colour information with kinematic information.

To bridge the gap between the BCJ double copy for amplitudes and the double copy for exact classical solutions, it will be useful to first review the story in the self-dual sector, as initiated by ref. [29]. Here a double copy can be set up between perturbative classical solutions to the equations of motion in self-dual Yang-Mills (SDYM) and self-dual gravity (SDG). Furthermore, due to the close relationship between perturbative classical solutions and tree-level amplitudes, this can be shown to be equivalent to the BCJ double copy for amplitudes. From here we will go on to review the first instance of an exact classical double copy that was found, the Kerr-Schild double copy of ref. [21]. Following a general discussion we will see a simple example of this formalism in action in the context of the Schwarzschild solution in general relativity. Finally, we will give an overview of the work that has been done to extend the exact classical double copy beyond Kerr-Schild solutions.

2.3.1 The double copy in the self-dual sector

The construction and interpretation of self-dual solutions in gauge theory and gravity will be discussed in Chapters 5 and 6. Here it is sufficient to state that self-dual gauge fields and metrics respectively satisfy

$$F_{\mu\nu} = \frac{i}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}, \qquad R_{\mu\nu\gamma\delta} = \frac{i}{2}\epsilon_{\mu\nu\rho\sigma}R^{\rho\sigma}{}_{\gamma\delta}, \tag{2.43}$$

where $F_{\mu\nu}$ is the field strength and $R_{\mu\nu\gamma\delta}$ is the the Riemann tensor. We work here in Minkowski spacetime and adopt the light-cone coordinates $x^\mu = \{u, v, w, \bar{w}\}$, where

$$u = t - z, \quad v = t + z, \quad w = x + iy, \quad \bar{w} = x - iy. \tag{2.44}$$

By now choosing the light-cone gauge $A_u = 0$ and considering a metric of the form

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}, \tag{2.45}$$

eqs. (2.43) imply that the non-zero components of A_μ and $h_{\mu\nu}$ are fixed in terms of scalar fields Ψ and ϕ respectively. Furthermore, eqs. (2.43) yield equations of motion for these scalars:

$$\Box\Psi + ig[\partial_u\Psi, \partial_w\Psi] = 0, \tag{2.46}$$

$$\Box\phi + \kappa\{\partial_u\phi, \partial_w\phi\} = 0. \tag{2.47}$$

Here $\Psi = \Psi^a T^a$ is valued in the Lie algebra of the gauge group, and we have introduced the Poisson bracket

$$\{f, g\} = (\partial_u f)(\partial_w g) - (\partial_w f)(\partial_u g). \tag{2.48}$$

The first equation of motion is often referred to simply as the self-dual Yang-Mills equation [56], while the second is the Plebanski equation [57]. Clearly these equations of motion are highly analogous, with the Lie bracket in SDYM replaced by a Poisson bracket in SDG. This relationship is formalised by colour kinematics duality. In momentum space, eqs. (2.46, 2.47) respectively take the form

$$\Psi^a(k) = -\frac{g}{2}\int đp_1 đp_2 \delta\!\!\!{}^-(p_1 + p_2 - k)\frac{X(p_1, p_2)}{k^2} f^{ab_1b_2}\Psi^{b_1}(p_1)\Psi^{b_2}(p_2), \tag{2.49}$$

$$\phi(k) = -\frac{\kappa}{2}\int đp_1 đp_2 \delta\!\!\!{}^-(p_1 + p_2 - k)\frac{X(p_1, p_2)^2}{k^2}\phi(p_1)\phi(p_2), \tag{2.50}$$

where

$$X(p_i, p_j) = p_{iw}p_{ju} - p_{iu}p_{jw}, \tag{2.51}$$

and we have introduced the following notation for convenience

$$đp \equiv \frac{d^4p}{(2\pi)^4}, \quad \delta\!\!\!{}^-(p) \equiv (2\pi)^4\delta^4(p). \tag{2.52}$$

We can now note that the non-linear term in the SDYM equations of motion consists of a colour structure constant f^{abc}, a propagator k^{-2}, and a factor which we can label

$$F_{p_ip_j}{}^{p_k} \equiv \delta\!\!\!{}^-(p_i + p_j - p_k)X(p_i, p_j). \tag{2.53}$$

These objects can be identified as the structure constants of an infinite-dimensional kinematic algebra. They are totally antisymmetric and satisfy the kinematic Jacobi identity

$$F_{p_1p_2}{}^q F_{p_3q}{}^k + F_{p_2p_3}{}^q F_{p_1q}{}^k + F_{p_3p_1}{}^q F_{p_2q}{}^k = 0, \tag{2.54}$$

where the contraction of indices is performed via an integral:

$$F_{p_1 q}{}^{k} F_{p_2 p_3}{}^{q} = \int đq \delta\!\!\!{}^{-}(p_1 + q - k) X(p_1, q) \delta\!\!\!{}^{-}(p_2 + p_3 - q) X(p_2, p_3) \tag{2.55}$$

$$= \delta\!\!\!{}^{-}(p_1 + p_2 + p_3 - k) X(p_1, p_2 + p_3) X(p_2, p_3). \tag{2.56}$$

Reference [29] identified the kinematic algebra as that of area-preserving diffeomorphisms in the (u, w) plane. The structure constants arise from the Poisson bracket evaluated on plane waves:

$$\{e^{-ik_1 \cdot x}, e^{-ik_2 \cdot x}\} = -X(k_1, k_2) e^{-i(k_1 + k_2) \cdot x}. \tag{2.57}$$

For a diffeomorphism in the (u, w) plane, the Poisson bracket is preserved if and only if it is area-preserving. The infinitesimal generators of such diffeomorphisms are

$$L_k = e^{-ik \cdot x}(-k_w \partial_u + k_u \partial_w), \tag{2.58}$$

and the Lie algebra is then

$$[L_{p_1}, L_{p_2}] = iX(p_1, p_2) L_{p_1 + p_2} = iF_{p_i p_j}{}^{p_k} L_{p_k}. \tag{2.59}$$

This is the algebra that is dual to the colour algebra under colour-kinematics duality. Perturbative solutions to the momentum-space equations of motion of eqs. (2.49, 2.50) with appropriate boundary conditions encode the tree-level amplitudes in the self-dual theories [29]. At each order in this perturbative expansion, one finds colour numerators built from contractions of the colour structure constants, and kinematic numerators built from analogous contractions of the kinematic structure constants of eq. (2.53). Whenever a set of three colour numerators satisfy the colour Jacobi identity, the associated kinematic numerators satisfy the kinematic Jacobi identity. Colour-kinematics duality is therefore manifest in the self-dual sectors.

With colour-kinematics duality manifest, it should be possible to double copy the perturbative SDYM solutions to obtain SDG solutions by replacing colour factors with kinematic factors. This is not quite as straightforward as making the replacement $f \to F$ in the terms of the perturbative expansion, as this would involve squaring delta functions. To identify the correct replacement, one can compute the n-point "off-shell amplitude" by functionally differentiating the nth-order term in the perturbative expansion with respect to the sources. The relevant colour and kinematic numerators can then be read-off, and the double copy takes the form of the standard BCJ procedure. In this way, the double copy for perturbative classical solutions in the self-dual sector is seen to be equivalent to the BCJ double copy for tree-level amplitudes.

The construction of perturbative classical solutions via the double copy has advanced a long way since this simple example. In ref. [58], perturbative solutions to equations of motion in full, non-self-dual Yang-Mills and general relativity were related via the double copy. Perturbative constructions also allow for the application of the double copy to dynamical classical problems such as black hole scattering and gravitational radiation. Following initial studies in refs. [59–61], this has rapidly flourished into a thriving field given the direct phenomenological applications. We defer to ref. [62] for a thorough review of these developments. This line of work has led to new methods for the extraction of classical physics from amplitudes [63–65], as well as the application of the double copy to worldline quantum field theory [66,67] and scattering on strong-field backgrounds [68].

2.3.2 The Kerr-Schild Double Copy

The fact that a double copy relation can be identified in perturbative classical solutions is in some sense unsurprising, given that these solutions encode tree-level amplitudes for which the BCJ double copy is well understood. Remarkably, however, it is also possible to interpret certain *exact* classical solutions in terms of the double copy paradigm. This was first outlined in ref. [21], the results of which we review here.

A family of exact solutions to Einstein's field equations are the Kerr-Schild metrics (for a detailed review see ref. [69]). In Kerr-Schild coordinates, the metric takes the form

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \kappa h_{\mu\nu}, \tag{2.60}$$

where $\bar{g}_{\mu\nu}$ is some background spacetime, κ is the gravitational coupling, and the deviation $h_{\mu\nu}$ from the background metric takes the form

$$h_{\mu\nu} = \phi k_\mu k_\nu. \tag{2.61}$$

Here ϕ is a scalar field and k is a vector defined to be both null and geodesic with respect to the background metric:

$$\bar{g}_{\mu\nu} k^\mu k^\nu = 0, \quad k^\mu \bar{\nabla}_\mu k_\nu = 0, \tag{2.62}$$

where $\bar{\nabla}_\mu$ is the covariant derivative associated with the background metric. This further implies that k is null and geodesic with respect to the full metric. Due to these properties the inverse metric takes the simple form

$$g^{\mu\nu} = \bar{g}^{\mu\nu} - \kappa\phi k^\mu k^\nu, \tag{2.63}$$

and it is therefore possible to raise and lower indices on k with either the background or full metric. Note that despite the similarity between eq. (2.60) and the linearised metric that characterises perturbative gravity in eq. (2.11), Kerr-Schild metrics are *exact* solutions to Einstein's equations. The field $h_{\mu\nu}$ thus describes some arbitrarily strong deviation from the background spacetime $\bar{g}_{\mu\nu}$. Despite this, it is customary to still refer to $h_{\mu\nu}$ as the graviton field.

The crucial feature of the Kerr-Schild form in eq. (2.61) is that the graviton field decomposes into an outer product of the vector k with itself. Remarkably, this form gives rise to a Ricci tensor that is linear in the graviton field and thus linearises the vacuum Einstein equations. Specifically, the components of the Ricci tensor are found to be

$$R^{\mu}{}_{\nu} = \bar{R}^{\mu}{}_{\nu} - \kappa h^{\mu}{}_{\rho}\bar{R}^{\rho}{}_{\nu} + \frac{\kappa}{2}\bar{\nabla}_{\rho}\left(\bar{\nabla}_{\nu}h^{\mu\rho} + \bar{\nabla}^{\mu}h^{\rho}{}_{\nu} - \bar{\nabla}^{\rho}h^{\mu}{}_{\nu}\right), \tag{2.64}$$

where $\bar{R}$ is the Ricci tensor corresponding to the background metric. It is important to note that the specific index placement in eq. (2.64) is required for linear dependence in $h_{\mu\nu}$.

In ref. [21] the particular case of stationary Kerr-Schild spacetimes with flat backgrounds, $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$, were considered. The background covariant derivative thus reduces to the flat-space derivative, and the full metric is time-independent such that

$$\partial_0\phi = 0, \quad \partial_0 k_{\mu} = 0. \tag{2.65}$$

We may furthermore set $k^0 = 1$ without loss of generality by absorbing all dynamics of the zeroth component into ϕ. With these conditions, the vacuum Einstein equations reduce to

$$R^{0}{}_{0} = \frac{1}{2}\partial^{i}\partial_{i}\phi = 0, \tag{2.66}$$

$$R^{i}{}_{0} = \frac{1}{2}\partial_{j}\left[\partial^{j}\left(\phi k^{i}\right) - \partial^{i}\left(\phi k^{j}\right)\right] = 0, \tag{2.67}$$

$$R^{i}{}_{j} = \frac{1}{2}\partial_{l}\left[\partial^{i}\left(\phi k^{l}k_{j}\right) + \partial_{j}\left(\phi k^{l}k^{i}\right) - \partial^{l}\left(\phi k^{i}k_{j}\right)\right] = 0, \tag{2.68}$$

where latin indices denote spacelike components. It is possible to interpret these equations in terms of the double copy. Consider pure Yang-Mills theory with equations of motion

$$\partial^{\mu}F^{a}_{\mu\nu} + gf^{abc}A^{\mu b}F^{c}_{\mu\nu} = 0. \tag{2.69}$$

As an ansatz for the gauge field let us take

$$A^{a}_{\mu} = \phi c^{a} k_{\mu}. \tag{2.70}$$

This can be obtained from the Kerr-Schild graviton in eq. (2.61) by replacing one of the Kerr-Schild vectors k_μ with a constant colour vector c^a. This yields a solution to the Yang-Mills equations, eq. (2.69), which linearise to give the abelian Maxwell equations due to the trivial colour dependence:

$$\partial^\mu F^a_{\mu\nu} = c^a \partial^\mu \left[\partial_\mu \left(\phi k_\nu\right) - \partial_\nu \left(\phi k_\mu\right)\right] = 0. \tag{2.71}$$

Writing the components of these equations explicitly, we find

$$\partial^\mu F^a_{\mu 0} = \partial^i \partial_i \phi = 0, \tag{2.72}$$
$$\partial^\mu F^a_{\mu i} = \partial^j \left[\partial_j \left(\phi k_i\right) - \partial_i \left(\phi k_j\right)\right] = 0, \tag{2.73}$$

which precisely correspond to eqs. (2.66) and (2.67) respectively. We have thus taken an exact solution to Einstein's equations, replaced kinematic information with colour information, and obtained an exact solution to the Yang-Mills equations. The analogy with the BCJ double copy for amplitudes has led this correspondence to be referred to as the Kerr-Schild double copy, with the gauge field in eq. (2.70) interpreted as the single copy of the graviton field in eq. (2.61).

To further extend the analogy with the perturbative double copy, we can take the zeroth copy of the gauge field in eq. (2.70) by replacing the vector k_μ with a second colour vector $\tilde{c}^{a'}$, to obtain

$$\Phi^{aa'} = \phi c^a \tilde{c}^{a'}. \tag{2.74}$$

This is a solution to a biadjoint scalar field theory, with equations of motion [21]

$$\partial^2 \Phi^{aa'} + y f^{abc} \tilde{f}^{a'b'c'} \Phi^{bb'} \Phi^{cc'} = 0, \tag{2.75}$$

where f^{abc} and $\tilde{f}^{a'b'c'}$ are the structure constants associated with two potentially different Lie algebras. Inserting the field in eq. (2.74) into eq. (2.75), we find that it abelianises the equations of motion to produce

$$\partial^2 \Phi^{aa'} = \partial^2 \phi = 0. \tag{2.76}$$

We may therefore interpret $\Phi^{aa'}$ as the zeroth copy of the gauge field in eq. (2.70). Note that the scalar field ϕ plays a role analogous to the propagators in the BCJ double copy, as it is present and unchanged in the biadjoint, gauge, and gravity solutions.

We have found a tower of exact solutions that are related by replacements between colour and kinematic information

$$\Phi^{aa'} = \phi c^a \tilde{c}^{a'}, \qquad A^a_\mu = \phi c^a k_\mu, \qquad h_{\mu\nu} = \phi k_\mu k_\nu \tag{2.77}$$

This mimics the situation represented in Figure 2.2. A natural question to ask is how seriously we should take this story for exact solutions as a double copy. In the BCJ double copy, the replacements between colour and kinematic numerators were justified by the fact that they obeyed identical algebraic constraints. That is to say, the double copy for amplitudes was made possible due to the presence of colour-kinematics duality. In the Kerr-Schild story it is unclear whether replacements between the colour and kinematic vectors, c^a and k_μ, can be similarly motivated. Furthermore, as noted in Section 2.2.3, taking two pure Yang-Mills theories as input into the BCJ double copy yields pure general relativity along with a dilaton and 2-form field, a theory known as $\mathcal{N} = 0$ supergravity. However, this seems to differ in the Kerr-Schild double copy, for which pure Yang-Mills maps to pure Einstein gravity. This is a consequence of the Kerr-Schild ansatz eq. (2.61). The 2-form field does not appear due to the symmetric nature of the ansatz, while the $k^2 = 0$ property sets the trace of the ansatz to zero, and thus the dilaton is not present [21].

These points suggest that the Kerr-Schild ansatz yields a rather special form of a double copy for classical solutions. Indeed more general methods have since been found that can be more closely identified with the double copy for amplitudes, and which reproduce the Kerr-Schild approach where overlap exists. We will review these developments at the end of this section, but first it will be useful to see the Kerr-Schild double copy in action in a simple example.

2.3.3 Example: the Schwarzschild solution

The Schwarzschild black hole is a static, asymptotically flat, spherically symmetric solution to the vacuum Einstein field equations. The Schwarzschild solution can be written in Kerr-Schild form, in which the metric takes the form

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{4G_N M}{r} k_\mu k_\nu, \tag{2.78}$$

where M is the positive, point-like mass of the source, r is the radial coordinate, and the Kerr-Schild vector k is

$$k^\mu = \left(1, \frac{x^i}{r}\right), \quad r^2 = \sum_{i=1}^{3} x^i x_i. \tag{2.79}$$

By comparing eq. (2.78) with eq. (2.60), we find the Kerr-Schild form of the graviton field to be

$$h_{\mu\nu} = \frac{\kappa}{2}\phi k_\mu k_\nu, \quad \phi = \frac{M}{4\pi r}. \tag{2.80}$$

To take the single copy of this solution, we make the replacements

$$\frac{\kappa}{2} \to g, \quad M \to c^a T^a, \quad k_\mu k_\nu \to k_\mu, \tag{2.81}$$

where T^a are the generators of the gauge group under consideration. This yields a non-abelian gauge field

$$A^a_\mu = \frac{gc^a}{4\pi r} k_\mu, \tag{2.82}$$

where $A_\mu = A^a_\mu T^a$. The replacements in eq. (2.81) make sense from the perspective of the double copy. In the first, the gravitational coupling is replaced by the gauge theory coupling, as in eq. (2.37) for the amplitude case. The second takes the gravitational charge, a mass, and replaces it with a colour charge. Finally, the third strips off one factor of the vector k_μ. As expected, the gauge field in eq. (2.82) linearises the Yang-Mills equations. It is a solution to the sourced Maxwell equations

$$\partial^\mu F^a_{\mu\nu} = j^a_\nu, \tag{2.83}$$

with a source

$$j^a_\mu = -gc^a u_\mu \delta^{(3)}(\mathbf{x}), \tag{2.84}$$

describing a static colour charge located at the origin with 4-velocity $u_\mu = (1, \mathbf{0})$. This is an interesting outcome. The Schwarzschild solution describes a point-like source at the origin. In taking the single copy, only the graviton and gauge fields have been taken into account, however the procedure has correctly identified the source required to generate the gauge field.

Ref. [21] continued to give a physical interpretation of the single copy gauge field in eq. (2.82). As the gauge field linearises the Yang-Mills equations, it is a solution to the abelian Maxwell equations. A gauge transformation therefore acts on the field as

$$A^a_\mu \to A'^a_\mu = A^a_\mu + \partial_\mu \Lambda^a(x), \tag{2.85}$$

for which we may choose

$$\Lambda^a = -\frac{gc^a}{4\pi} \log\left(\frac{r}{r_0}\right), \tag{2.86}$$

where r_0 is an arbitrary length scale introduced to ensure the logarithm argument is dimensionless. This transformation gives rise to a second gauge field

$$A'^a_\mu = \frac{gc^a}{4\pi r} u_\mu, \tag{2.87}$$

where $u_\mu = (1, \mathbf{0})$. This is the Coulomb potential for a static point colour charge located at the origin; the most general time-independent spherically symmetric solution in linearised gauge theory.

It is important to note that this process was only possible due to the Kerr-Schild form of the metric, and the resultant linearisation of Einstein's equations. Unlike BCJ duality, in which the amplitudes identified on either side of the correspondence are gauge invariant objects, for the classical double copy to work the solutions must be expressed within a particular gauge. Indeed, the two gauge equivalent solutions in eqs. (2.82) and (2.87) give rise to different gravity solutions upon taking the double copy. There is at present no way to circumvent this issue. Thus, in finding new examples of the classical doubly copy, it is easier to begin with the gravity solution and single copy to find the corresponding gauge field.

2.3.4 Beyond Kerr-Schild solutions

The Kerr-Schild double copy reviewed here provided the first example of an exact double copy. Since this point, many other exact classical solutions have been furnished with a double copy interpretation. With the Kerr-Schild double copy in hand, it should in principle be possible to identify the single copy of all vacuum metrics that admit a Kerr-Schild or multi-Kerr-Schild form. Indeed, in ref. [70], the full family of Kerr-Taub-NUT metrics in vacuum gravity were single copied. Of particular interest here is Taub-NUT spacetime which single copies to an electromagnetic dyon. This will play an important role in chapter 3. The Kerr-Schild double copy has further been generalised to time-dependent solutions [71], as well as to cases where the Kerr-Schild graviton is a deviation around a curved, rather than Minkowski, background [72]. The most general double copy for pure Yang-Mills theory is NS-NS gravity. This fact was incorporated into the Kerr-Schild double copy in ref. [73], in which it was found that the double copy of a point charge is the Janis-Newman-Winicour (JNW) metric. In ref. [21], a Kerr-Schild-like approach to self-dual solutions was given, in which the graviton is written in terms of certain differential operators (we review this story in section 5.1.3). This construction led to a single copy of the Eguchi-Hanson metric and a first look at the role of topology in the double copy [74]. There also exist extensions of the Kerr-Schild double copy formalism to double field theory and supergravity [75,76], exceptional field theory [77], spacetimes with non-zero cosmological constants [78], and Kaluza-Klein theory [79].

Despite the incredible amount of progress that has been made with the Kerr-Schild double copy formalism, the Kerr-Schild family of metrics are highly special. One might then ask if an exact double copy can be developed for more general gravitational solutions. In ref. [22], a procedure for taking the single copy of all vacuum type-D and certain type-N metrics was developed, a formalism known as the *Weyl double copy*. Here "type" refers to the algebraic classification of spacetimes introduced by Petrov (see e.g. ref. [69] for a review). The Weyl double copy makes use of the spinorial formulation of gauge theory and gravity, and provides a double copy in terms of the Weyl and Maxwell spinors in these theories. For certain vacuum metrics, the Weyl spinor $\Psi_{\alpha\beta\gamma\delta}$ decomposes as [22]

$$\Psi_{\alpha\beta\gamma\delta} = \frac{\Phi_{(\alpha\beta}\Phi_{\gamma\delta)}}{S}, \tag{2.88}$$

where $\Phi_{\alpha\beta}$ is the Maxwell spinor, satisfying the vacuum Maxwell equations in flat spacetime, and S is a scalar field which satisfies the flat spacetime wave equation. All type-D and type-N solutions posses a Weyl spinor which decomposes into the specific spinorial form of eq. (2.88). It can then be shown that the fields on the right-hand side correspond to Maxwell spinors for all type-D solutions [22] and for non-twisting type-N solutions [80]. The Weyl double copy is particularly nice in that it can be shown to provide an explicit connection between the double copy for amplitudes and the double copy for exact classical solutions [81, 82].

Type-D solutions contain all metrics which admit a Kerr-Schild form, but also additional solutions which do not, such as the C-metric. The Weyl double copy thus generalises the Kerr-Schild double copy to a larger class of exact spacetimes, and agrees with its results where there is overlap. The Weyl double copy has also been derived from twistor space, a formalism known as the *twistor double copy* [23–25]. This provides a double copy interpretation of type-III solutions as well as the multipole expansion of vacuum type-D solutions [83]. Alternative exact double copy constructions exist, such as the convolutional approach of refs. [84–87]. As in the Kerr-Schild and Weyl double copies, this approach applies at the linearised level, however it has led to insights into the double copy structure of non-linear interacting Lagrangians [53–55].

Chapter 3

Topology and the double copy

In its traditional form the double copy is a relation between *local* quantities in different theories. In the case of the BCJ double copy, locality is built into scattering amplitudes by definition. Similarly, the classical double copy for exact solutions relates fields defined at the same spacetime point and is thus a local statement. In extending the remit of the double copy it is therefore interesting to ask whether it can be generalised in some form to global information, such as non-trivial topology. This is a particularly compelling question from the perspective of the classical double copy. Topology sometimes plays an integral role in the construction of exact solutions to equations of motion, such as in the classification of possible solutions. We can therefore ask whether it is possible to identify a double copy between exact solutions each of which exhibits a topological characterisation that matches up on either side of the correspondence.

This question was first addressed in ref. [74]. In this work, the classical double copy for self-dual solutions was studied, building on the Kerr-Schild-like construction for such solutions first introduced in ref. [21]. This led to the identification of a single copy for the Eguchi-Hanson instanton in pure general relativity. At face value, this set-up appears to be the perfect situation to investigate the role of topology in the double copy, as instantons are inherently topological constructions. As demonstrated in ref. [74], the Eguchi-Hanson spacetime is in general a topologically trivial solution. However, by restricting one of the coordinates on the spacetime, one can induce a non-trivial topology. The single copy of the Eguchi-Hanson metric was found to be an abelian-like gauge field, as in all previous cases of the classical double copy. By definition, such fields are topologically trivial, which is consistent with its gravitational counterpart. However, it is then natural to ask whether the coordinate restriction considered for the gravitational solution likewise induces a non-trivial topology for the single copy field. Ref. [74] found that this was not the case, such that the gauge field remained topologically trivial after restricting the coordinate. In some sense, this is not a surprising result. The coordinate restriction in gravity plays the important role of removing a singularity in the field, which in turn gives rise to the topological character. There is, however, no analogous singularity in gauge theory, such that the single copy solution remains topologically trivial even after the coordinate restriction. Thus, in this case, it appears that there is no way in which a counterpart to the non-trivial topology of a solution on one side of the double copy can be identified on the other.

In this chapter we return to this question, now in the context of a different class of topologically non-trivial solutions: magnetic monopoles. Monopoles are exact gauge theory solutions which arise from the introduction of sources in the magnetic field. While such fields may at first seem prohibited by classical electromagnetism, we will review in the following section an argument proposed by Dirac to circumvent this issue. From the perspective of the double copy, monopoles appear in the single copy of the Taub-NUT metric. In ref. [70], electromagnetic dyons were identified as the single copy of the Taub-NUT solution, where the electric and magnetic charges are dual to the mass and NUT charges in the gravitational solution respectively. Magnetic monopoles then correspond to the single copy of the zero mass limit of the Taub-NUT metric, known as the pure NUT solution.

In addition to their topological nature, monopole solutions are interesting from the perspective of the double copy due to the insights they provide into the non-linear nature of the correspondence at the level of exact classical solutions. In almost all examples of the classical double copy, the gauge theory solution is an abelian-like field, such that it linearises the full Yang-Mills equations. However, as colour information is removed when taking the double copy, one might expect that solutions with different gauge groups should map to the same gravitational solution. This was indeed the situation found in ref. [88]. Here it was shown that an abelian-like Dirac monopole and the non-abelian Wu-Yang monopole are related by a singular gauge transformation [89,90].[1] By an abelian-like Dirac monopole, we mean a non-abelian dressing of the Dirac monopole, whose colour structure linearises the Yang-Mills equations. As the Dirac monopole is related to the pure-NUT solution via the double copy, one can therefore consider *either* a fully non-abelian or an abelian-like gauge field as double copying to the same gravitational solution. This situation, however, appears to pose a problem in attempting to map topological information under the double copy. Topological invariants in gauge

[1] Note that the Wu-Yang monopole is not the same as the Wu-Yang construction of the Dirac monopole considered later in this chapter. The former is a fully non-abelian monopole solution, while the latter is a construction of the abelian Dirac monopole in which the wire-like singularities are removed.

theory usually rely on the particular gauge group under consideration, whereas the story outlined here suggests that this is irrelevant from the perspective of the double copy.

In this chapter we will use magnetic monopole solutions as a playground to investigate both the non-linear nature of the exact double copy and the role of topology in the correspondence. We will begin with a brief review of monopole solutions, Taub-NUT spacetime, and the double copy between them. Following this we will discuss the general construction of non-abelian monopole gauge fields, from which we will see that it is always possible to write the non-abelian field as a dressed abelian-like solution, such that the double copy may be simply performed. This will lead naturally into a discussion of the topology of monopole solutions, both abelian and non-abelian, and the Taub-NUT solution. We will see that by recasting the non-trivial topology of all of these solutions in terms of a patching condition between fields defined on different domains, we obtain a topological characterisation which matches up on either side of the double copy correspondence, regardless of the choice of gauge group. This will therefore correspond to an exact classical solution whose local and global properties follow the same double copy construction.

3.1 Monopoles and the double copy

3.1.1 Magnetic monopoles

In this section we introduce the basics of magnetic monopoles. Excellent reviews of this material can be found in refs. [91, 92]. Consider the source free Maxwell equations in 3+1 dimensions[2]

$$\partial_\mu F^{\mu\nu} = 0, \qquad \partial_\mu \tilde{F}^{\mu\nu} = 0, \tag{3.1}$$

where $F_{\mu\nu}$ are the abelian field strength components

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{3.2}$$

and $\tilde{F}_{\mu\nu}$ are the components of the dual field strength $*F$:

$$*F = \tilde{F}_{\mu\nu} dx^\mu \wedge dx^\nu, \qquad \tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}. \tag{3.3}$$

[2]Note that we use a $(+---)$ metric signature in this chapter alone, so as to maintain consistency with the literature.

The vacuum Maxwell equations are invariant under Lorentz and conformal transformations, as well as electromagnetic duality

$$F \to *F, \qquad *F \to -F, \tag{3.4}$$

where the minus arises due to the identity $*^2 = -1$. At the level of the vector fields $F_{0i} = E_i$ and $F_{ij} = \epsilon_{ijk} B_k$, this duality transformation is

$$\boldsymbol{E} \to \boldsymbol{B}, \qquad \boldsymbol{B} \to -\boldsymbol{E}, \tag{3.5}$$

and thus constitutes a symmetry between the electric $\boldsymbol{E}$ and magnetic $\boldsymbol{B}$ fields. The sourced Maxwell equations modify eqs. (3.1) to include electric charges but no magnetic analogues

$$\partial_\mu F^{\mu\nu} = j_{\mathrm{e}}^\nu, \qquad \partial_\mu \tilde{F}^{\mu\nu} = 0, \tag{3.6}$$

thereby breaking the symmetry under electromagnetic duality. It is therefore tempting to conjecture the existence of poles in the magnetic field, such that the equations governing classical electromagnetism maintain their full symmetries. At first sight, such a modification seems problematic. The fact that the magnetic field is divergenceless in standard classical electromagnetism implies that it can be written globally as the curl of a vector potential, thereby permitting the standard definition of the abelian field strength in terms of the gauge field A_μ, as in eq. (3.2). The introduction of magnetic sources j_{m}^ν, such that

$$\partial_\mu \tilde{F}^{\mu\nu} = j_{\mathrm{m}}^\nu, \tag{3.7}$$

prohibits the introduction of the gauge potential. There is, however, a loophole to this argument. Provided the magnetic sources are point-like, it is possible to define the gauge field in regions for which $j_{\mathrm{m}}^\nu = 0$. Generically, the topology of such regions will be non-trivial, and thus it may not be possible to define a non-singular gauge field everywhere. However, we can obtain a sensible theory provided a given gauge field is locally well-defined, and related to those defined elsewhere via a gauge transformation in the regions where their domains overlap. Dirac proposed the existence of magnetic sources defined in this way, referred to as *magnetic monopoles* or simply *monopoles*. Specifically abelian monopoles are referred to as *Dirac monopoles*.

The Dirac monopole

Magnetic monopoles are point-like sources of the magnetic field. A monopole located at the origin in $\mathbb{R}^3$ will generate a Coulomb magnetic field

$$\boldsymbol{B} = \frac{Q_M}{4\pi r^2}\hat{\boldsymbol{e}}_r. \tag{3.8}$$

Here Q_M is the magnetic charge and the factor of 4π is included in analogy to the standard definition of the electric Coulomb field

$$\boldsymbol{E} = \frac{Q_E}{4\pi r^2}\hat{\boldsymbol{e}}_r. \tag{3.9}$$

We wish to construct a gauge field which, upon taking the curl, gives rise to the magnetic field in eq. (3.8). The field strength defined from this gauge field should be continuous and single-valued. The most straightforward construction of such a gauge field was given by Wu and Yang in ref. [93], and is referred to as the *Wu-Yang construction* of the Dirac monopole. The approach is to define two separate gauge fields on a S^2 sphere centred on the monopole. Writing the magnetic charge of the monopole as $Q_M = 4\pi\tilde{g}$ for notational convenience, consider the following two vector potentials

$$A_i^{\mathrm{N}} = -\frac{\tilde{g}\epsilon_{ij3}x^j}{r(z+r)}, \quad A_i^{\mathrm{S}} = -\frac{\tilde{g}\epsilon_{ij3}x^j}{r(z-r)}. \tag{3.10}$$

In addition to the point-like singularities at the origin $r = 0$, both of these solutions feature line-like singularities. The A_i^N field is singular everywhere along the negative z-axis, while A_i^S is singular along the positive z-axis. These singularities are known as *Dirac strings*. Away from the singularities the fields in eq. (3.10) are locally well-defined and their curl generates the magnetic field in eq. (3.8). A_i^N and A_i^S can thus be taken to define a monopole potential in the the northern and southern hemispheres of the two-sphere respectively, hence the labels N and S. Figure 3.1 provides a pictorial representation of this setup.

The location of the Dirac strings seems somewhat arbitrary. Indeed, note that the difference between the northern and southern fields is a gradient:

$$\begin{aligned} A_i^N - A_i^S &= -\frac{\tilde{g}\epsilon_{ij3}x^j}{r}\left(\frac{1}{z+r} - \frac{1}{z-r}\right) \\ &= 2\tilde{g}\partial_i\left[\arctan\left(\frac{y}{x}\right)\right] \\ &= 2\tilde{g}\partial_i\phi, \end{aligned} \tag{3.11}$$

where in the final equality we have identified the azimuthal angle on the two-sphere $\phi =$

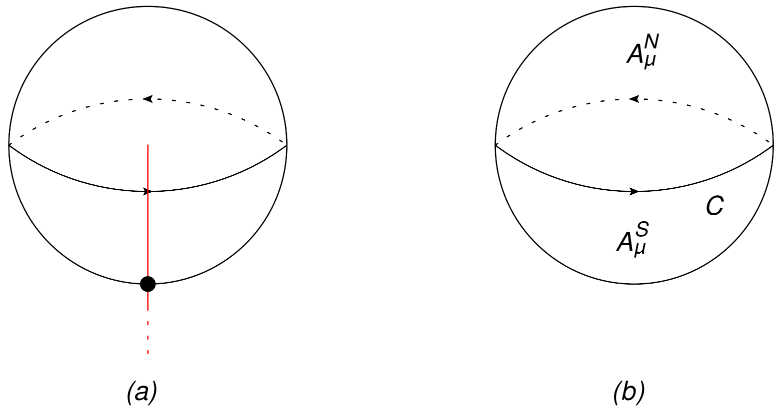


Figure 3.1: (a) The northern Dirac monopole field $A_i^{\rm N}$ in eq. (3.10). The red line represents the Dirac string singularity, which is aligned with the negative z-axis and intersects the sphere at the south pole. (b) The Wu-Yang construction of the Dirac monopole. Two separate fields are defined, that are non-singular within their respective domains. On the region of overlap, which we take to be the equator, they are related by a gauge transformation.

$\arctan(y/x)$. This difference therefore corresponds to an abelian gauge transformation, that is singular on the location of the initial and final strings. In this way, via an appropriate gauge transformation, the Dirac string of an abelian monopole field can be oriented to coincide with an arbitrary curve running from the origin out to infinity.

In the following, it will be useful to rewrite the vector potentials of eqs. (3.10) in spherical polar coordinates and to include them as components of a four-dimensional gauge field. Consider the case for which there is zero electric potential, such that $A_0 = 0$. Then only the ϕ-components of the monopole gauge fields are non-zero:

$$A_\phi^{\rm N} = -\tilde{g}\left(\cos\theta - 1\right), \quad A_\phi^{\rm S} = -\tilde{g}\left(\cos\theta + 1\right). \tag{3.12}$$

These fields are related by the abelian gauge transformation

$$A_\mu^{\rm N} = A_\mu^{\rm S} + \frac{i}{g} S(\phi)\partial_\mu S^{-1}(\phi), \tag{3.13}$$

where g is the coupling and $S(\phi)$ is an element of the U(1) gauge group,

$$S(\phi) = e^{2ig\tilde{g}\phi}. \tag{3.14}$$

We thus have two gauge fields that are non-singular everywhere within their own hemispheres. We can take their region of overlap to be the equator, $\theta = \pi/2$, on which they are equal up to the gauge transformation of eq. (3.13). When taken together, the two gauge fields in eq. (3.12) constitute a well-defined monopole solution away from the origin. To check that the fields defined act as expected, we can calculate the magnetic flux over the two-sphere at infinity

$$\Phi_B = \iint_{S^2} d\Sigma^{\mu\nu} F_{\mu\nu}, \tag{3.15}$$

where $d\Sigma^{\mu\nu}$ is the area element on the surface. Separating this into separate contributions from each hemisphere and applying Stokes' theorem gives

$$\Phi_B = \oint_C dx^\mu \left(A^{\mathrm{N}}_\mu - A^{\mathrm{S}}_\mu\right) = 4\pi\tilde{g} = Q_{\mathrm{M}}. \tag{3.16}$$

This is the desired result, with Q_{M} denoting the magnetic charge within the sphere.

Dirac's quantisation condition

The presence of magnetic monopoles leads to some surprising conclusions when we attempt to develop a consistent quantum theory involving magnetic sources. Suppose that an electrically charged particle traverses a closed curve C in the presence of a monopole, such that the area enclosed by C feels some non-zero magnetic flux. In the quantum theory, the electrically charged particle's wavefunction will pick up a phase after completing this loop, such that

$$\psi \to U_C \psi, \tag{3.17}$$

where U_C is the Aharonov-Bohm phase factor

$$U_C = \exp\left(ig \int_C d\boldsymbol{x} \cdot \boldsymbol{A}\right). \tag{3.18}$$

Let us now suppose that the curve C encloses a Dirac string. As this singularity is a gauge dependent artifact, it should be unobservable. Thus, in the limit in which the curve is contracted to infinitesimal loop around the string, the phase factor obtained in the presence of the monopole should match that obtained when the monopole is absent,

$$U_C\big|_{\tilde{g}=0} = U_C\big|_{\tilde{g}\neq 0} \tag{3.19}$$

Consider the northern hemisphere monopole field in eq. (3.12), for which the singularity lies along the negative z-axis. An infinitesimal loop around the string thus corresponds to the limit $\theta \to \pi$. In the case in which there is no monopole, for this infinitesimal loop we simply have $U_C|_{\tilde{g}=0} = 1$. With the monopole present we can perform the integral to find

$$U_C\big|_{\tilde{g}\neq 0} = \exp\left(ig\int d\phi A_\phi^{\rm N}\right)$$
$$= \exp\left(-2\pi i g\tilde{g}(\cos\theta - 1)\right) \tag{3.20}$$

Equation (3.19), in the $\theta \to \pi$ limit, therefore translates to

$$1 = e^{4\pi i g\tilde{g}}, \tag{3.21}$$

such that

$$g\tilde{g} = \frac{n}{2}, \quad n \in \mathbb{Z}. \tag{3.22}$$

This result is known as the *Dirac quantisation condition.* Note that we could have similarly found this result by requiring that the U(1) element in eq. (3.14) defining the gauge transformation be single valued, such that $S(0) = S(2\pi)$.

In deriving the Dirac quantisation condition we have made no assumptions about the electric or magnetic charges involved, and thus the quantisation condition must hold for all g and $\tilde{g}$. This is only possible if

$$g = \alpha g_{\rm min}, \quad \tilde{g} = \beta \tilde{g}_{\rm min}; \qquad \alpha, \beta \in \mathbb{Z}, \tag{3.23}$$

and

$$g_{\rm min}\tilde{g}_{\rm min} = \frac{1}{2}. \tag{3.24}$$

Thus, by requiring that the gauge-dependent monopole singularities are unobservable, we have stumbled across the quantisation of electric charge. In the modern usage, it is not thought that the Dirac quantisation condition is responsible for the quantised electric charge that is observed in nature, as the Standard Model already provides an explanation for this. Instead it can be said that *because* electric charge is quantised, the quantisation condition of eq. (3.22) permits the existence of magnetic monopoles.

More exotic monopoles

In this section we have given a brief outline of the description and physical consequences of sources in the magnetic field.[3] In particular, we have considered the simplest magnetic monopoles, those that are abelian and singular at the origin. By relaxing both of these conditions, a far more diverse catalogue of monopoles can be formulated.

[3] For a recent discussion of the experimental status of magnetic monopoles, see ref. [94].

Non-abelian monopoles can be constructed either by embedding the U(1) gauge group of the Dirac monopole in a larger gauge group, or via a particular singular gauge transformation. This second approach yields the so-called Wu-Yang monopole, not to be confused with the Wu-Yang construction of abelian monopoles described earlier. We will consider the case of non-abelian monopoles in detail in the coming sections.

The construction of finite energy monopoles that are non-singular at the origin is also possible, and leads to the *'t Hooft Polyakov monopole*. 't Hooft Polyakov monopoles are non-abelian solutions that necessarily appear when Yang-Mills is coupled to a Higgs field, which by the spontaneous symmetry breaking mechanism causes the gauge group to reduce to a smaller subgroup. The existence of these solutions has interesting consequences for the development of grand unified theories [92]. We will not consider 't Hooft Polyakov monopoles in this work. At present, no double copy interpretation of these solutions exists, and whether such a correspondence with a gravitational solution can be set up remains an open question.

3.1.2 The Taub-NUT metric

The Taub-NUT solution [95, 96] is a stationary, axisymmetric solution to the vacuum Einstein equations (for a modern review, see ref. [97]). It has the important property that it is not asymptotically flat, but instead has a non-zero rotational character at spatial infinity. It is therefore not a particularly physical solution, but nevertheless it continues to play an important role in modern theoretical physics. The metric is classified by two parameters: a Schwarzschild-like mass term and the so-called NUT charge, which generates the rotational component of the field.

In spherical coordinates, the Taub-NUT metric is

$$ds^2 = f(r)\left[dt + 2N\cos\theta d\phi\right]^2 - f^{-1}(r)dr^2 - \left(r^2 + N^2\right)d\Omega^2, \tag{3.25}$$

where

$$d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2 \tag{3.26}$$

is the metric on the two-sphere at constant radius, and the radial function $f(r)$ is given by

$$f(r) = \frac{(r-r_+)(r-r_-)}{r^2+N^2}, \quad r_\pm = M \pm \sqrt{M^2+N^2}. \tag{3.27}$$

M and N are the mass and NUT charge respectively. Note that for $N \to 0$, the Taub-NUT solution reduces to the Schwarzschild solution. In the following, we will be interested in the converse limit known as the *pure NUT* form of the metric, for which we take the mass to vanish, $M \to 0$. The pure NUT metric is then given by

$$ds^2 = \frac{r^2 - (\kappa N)^2}{r^2 + (\kappa N)^2} \left[dt + 2\kappa N \cos\theta d\phi\right]^2 - \left(r^2 + (\kappa N)^2\right) \left[\frac{dr^2}{r^2 - (\kappa N)^2} + d\Omega^2\right], \quad (3.28)$$

where the rescaling $N \to \kappa N$ has been made for later convenience. Both the Taub-NUT and pure NUT metric have coordinate singularities at $\theta = 0$ and $\theta = \pi$, where they cannot be inverted. These singularities extend along the entire axes defined by $\theta = 0$ and $\theta = \pi$, which can be taken to be the positive and negative z-axes respectively. They are known as *Misner strings*, first discussed in ref. [98], and are clear analogues of the Dirac strings found in magnetic monopole gauge fields. The similarities between the Taub-NUT metric and monopole gauge fields continue in the construction of a metric that is free of these wire-like singularities. This can be done by dividing spatial slices into two hemispheres, with the northern hemisphere defined by $\theta \in [0, \pi/2]$ and the southern by $\theta \in [\pi/2, \pi]$. Within each hemisphere we may perform the following transformations of the time coordinate:

$$\theta \in [0, \pi/2]: \quad t \to t_{\mathrm{N}} = t + 2\kappa N \phi, \quad (3.29)$$

$$\theta \in [\pi/2, \pi]: \quad t \to t_{\mathrm{S}} = t - 2\kappa N \phi, \quad (3.30)$$

where the N and S labels correspond to the northern and southern hemispheres respectively. Performing these transformations in eq. (3.28), we obtain two forms for the pure NUT metric:

$$ds^2_{\mathrm{N,S}} = \frac{r^2 - (\kappa N)^2}{r^2 + (\kappa N)^2} \left[dt + 2\kappa N \left(\cos\theta \mp 1\right) d\phi\right]^2 - \left(r^2 + (\kappa N)^2\right) \left[\frac{dr^2}{r^2 - (\kappa N)^2} + d\Omega^2\right], \quad (3.31)$$

where the northern (southern) label corresponds to the upper (lower) sign in $\mp$. The northern metric is singular at $\theta = \pi$ and is therefore non-singular everywhere in the northern hemisphere. Similarly, the singularity in the southern metric now coincides with $\theta = 0$ and it is thus non-singular everywhere in the southern hemisphere. Furthermore, in the overlap region, which corresponds to the equator at $\theta = \pi/2$, the northern and southern time coordinates are related by

$$t_{\mathrm{N}} = t_{\mathrm{S}} + 4\kappa N \phi. \quad (3.32)$$

The coordinate ϕ is compact with period 2π, such that

$$\begin{aligned} t_{\mathrm{N}} &= t_{\mathrm{S}} + 4\kappa N(\phi + 2\pi) \\ &= t_{\mathrm{N}} + n t_0, \qquad n \in \mathbb{Z} \end{aligned} \quad (3.33)$$

and similarly for t_{S}. We therefore find that the time coordinates in each hemisphere are compact with period

$$t_0 = 8\pi\kappa N_0, \tag{3.34}$$

where N_0 can be interpreted as a minimum unit of NUT charge. The full NUT charge that appears in the metric is then defined by

$$N = nN_0, \quad n \in \mathbb{Z}. \tag{3.35}$$

The construction of Taub-NUT spacetime in terms of two separate fields therefore enforces the quantisation of NUT charge.

This process of defining a metric for Taub-NUT spacetime that is free of Misner strings is highly analogous to the process of defining a gauge field for the Dirac monopole that is free of Dirac strings, as was outlined in the previous section. In both cases we introduce two distinct forms for the metric/gauge field. These each contain a wire-like singularity intersecting one of the poles of a spatial two-sphere, but are finite everywhere else. They can thus be taken to be well-defined on the northern and southern hemispheres of the two-sphere respectively. On the region of overlap, which we can take to be the equator, the two forms of the metric/gauge field are related by a transformation. For the monopole gauge fields this corresponds to an abelian gauge transformation, while for the Taub-NUT metric it is a coordinate transformation. Furthermore, the Dirac quantisation condition finds a natural partner in the quantisation of NUT charge arising due to the periodicity of the time coordinate. Given this discussion, one can ask: are these similarities merely a coincidence or are they indicative of some deeper correspondence between these two field configurations? The double copy provides a possible answer to this question.

3.1.3 The double copy of an abelian monopole

It has long been known that the Taub-NUT solution in general relativity and magnetic monopole solutions in abelian gauge theory are highly analogous (see e.g. ref. [97] for a discussion). In fact, the extent of the similarities between the two solutions has led to the NUT charge being referred to as a gravitational magnetic charge. One source of this identification arises from the study of the Euclidean Taub-NUT metric. Here, via the compactification of the Euclidean time coordinate, one finds a magnetically charged black hole. For this reason, the Euclidean Taub-NUT metric is sometimes known as a Kaluza-Klein monopole.

Here we will outline a complementary story, in which the Taub-NUT solution is identified as the double copy of the abelian magnetic monopole. That such a double copy is possible is suggested already by the form of the Taub-NUT metric and the known single copy properties of the Schwarzschild solution. Recall that the $N \to 0$ limit of Taub-NUT reduces the solution to the Schwarzschild metric. This can be interpreted as the double copy of a Coulomb electric charge, as was discussed in section 2.3.3, where the Schwarzschild mass M is dual, in the sense of the double copy, to the electric charge. As Taub-NUT is equivalent to Schwarzschild but with an additional NUT charge N turned on, and as N plays the role of a sort of gravitational magnetic charge, it is natural to assume that the single copy of the Taub-NUT solution should be a dyon, a point-like source with both electric and magnetic charge. While this identification is intuitive, it does not tell us how to single copy Taub-NUT in practice. A prescription for this was given in ref. [70].

Underlying the ability to single copy the Schwarzschild solution was the fact that it can be written in Kerr-Schild form. When written in so-called Plebanksi coordinates [99], the Taub-NUT solution exhibits a *double Kerr-Schild form* [100]. That is, the metric takes the form

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \kappa h_{\mu\nu}, \tag{3.36}$$

where $\bar{g}_{\mu\nu}$ is a background metric and the graviton $h_{\mu\nu}$ decomposes into[4]

$$h_{\mu\nu} = \phi M k_\mu k_\nu + \psi N l_\mu l_\nu. \tag{3.37}$$

Here ϕ, ψ are scalars and M, N are the mass and NUT charge respectively. The vectors k_μ and l_μ both satisfy the null and geodesic conditions of Kerr-Schild vectors

$$\bar{g}_{\mu\nu} k^\mu k^\nu = \bar{g}_{\mu\nu} l^\mu l^\nu = 0, \quad k^\mu \bar{\nabla}_\mu k_\nu = l^\mu \bar{\nabla}_\mu l_\nu = 0, \tag{3.38}$$

as well as a mutual orthogonality condition

$$\bar{g}_{\mu\nu} k^\mu l^\nu = 0. \tag{3.39}$$

This form for the metric is thus a simple extension of the standard Kerr-Schild coordinates. In general a double Kerr-Schild form will not linearise the Einstein equations. However, for the Taub-NUT metric in Plebanski coordinates the non-linear part of the Ricci tensor vanishes, and thus linearisation is achieved [100].

It is possible to take the single copy of the Taub-NUT graviton in eq. (3.37). To do so we work term-by-term in the two terms that make up the double Kerr-Schild form. This can be thought of as analogous to how the BCJ double copy for amplitudes applies separately for terms involving distinct propagators. Thus, we make replacements in

[4] For the explicit form of the Taub-NUT metric in double Kerr-Schild coordinates see ref. [70].

the first term of eq. (3.37) corresponding to taking the single copy of the Schwarzschild solution

$$M \to c^a T^a, \quad k_\mu k_\nu \to k_\mu, \tag{3.40}$$

along with analogous replacements in the second term

$$N \to \tilde{c}^a T^a, \quad l_\mu l_\nu \to l_\mu. \tag{3.41}$$

The result is the gauge field

$$A_\mu = \phi c^a T^a k_\mu + \psi \tilde{c}^a T^a l_\mu, \tag{3.42}$$

which solves the linearised Yang-Mills equations. Ref. [70] went on to interpret this gauge field as that corresponding to a dyon, a point-like source possessing both electric and magnetic charge. The electric charge corresponds to the first term in eq. (3.42), while the magnetic charge corresponds to the second. Thus, as expected, the mass and NUT charge in gravity are respectively dual to the electric and magnetic charges in gauge theory under the double copy.

3.1.4 Abelian, abelian-like, and non-abelian single copies

At this point, it will be useful to pause and discuss the sense in which single copy gauge fields are non-abelian. A general gauge field obtained via the Kerr-Schild double copy will take the form

$$A_\mu = \phi c^a T^a k_\mu. \tag{3.43}$$

This is a non-abelian gauge field and can be defined for any gauge group. It is therefore a solution to the non-linear Yang-Mills equations, however it is a highly special solution in that it linearises these equations of motion. This makes perfect sense from the perspective of the Kerr-Schild double copy. The linearisation of the Yang-Mills equations corresponds simply to the fact that the gauge field has been obtained by single copying a metric that takes a Kerr-Schild form, and thus linearises the Einstein equations.

Gauge fields obtained via the Kerr-Schild single copy are therefore *abelian-like* solutions, satisfying the Maxwell equations. As such, it will be instructive in the following to rewrite eq. (3.43) as

$$A_\mu = c^a T^a A_\mu^{\text{abel.}} \tag{3.44}$$

where $A_\mu^{\text{abel.}}$ is a truly abelian gauge field. The important point here is that if $A_\mu^{\text{abel.}}$ has a known Kerr-Schild double copy, any non-abelian solution of this form, in which the colour structure factorises completely, can also be straightforwardly double copied.

One simply removes the colour charge, and double copies the gauge field as if it were abelian. In this way, it appears that at least in some cases both abelian and non-abelian gauge fields double copy to the same solution in gravity.

This discussion is relevant in the case of the double copy between magnetic monopoles and Taub-NUT spacetime for the following reason. In the case of the pure NUT solution, for which $M \to 0$, the single copy corresponds to a magnetic monopole and can be written in the form

$$A_\mu = \tilde{c}^a T^a A^{\mathrm{D}}_\mu , \tag{3.45}$$

where A^{D}_μ corresponds to a Dirac monopole

$$A^{\mathrm{D}}_\phi = -\tilde{g}\,(\cos\theta - 1)\,. \tag{3.46}$$

Here we have chosen the northern hemisphere gauge field in eq. (3.12), however the southern field works just as well. Note that this takes exactly the form of eq. (3.44) with the Dirac monopole playing the role of the abelian gauge field. This is unsurprising as it was obtained via the Kerr-Schild single copy. In ref. [70] it was not explicitly checked whether eq. (3.45) corresponds to a genuine non-abelian monopole solution. In the following, however, we will see that this is the case, and therefore singular monopoles lie within the class of solutions for which the abelian and non-abelian double copies align.

3.2 The double copy of a non-abelian monopole

In this section we will examine the structure of point-like non-abelian monopoles, that are singular at the origin. Much of the literature on non-abelian monopoles focuses on non-singular solutions, which arise in spontaneously broken theories with additional scalar fields (for reviews see e.g. ref. [92]). It is not yet known how to double copy such solutions, and so here we focus on the singular case. The key to this analysis will be the ability to find a gauge in which the non-abelian monopole gauge fields take a form in which the colour structure completely factorises. The fact that this is possible is not new information, however its interpretation in terms of the double copy is novel. Furthermore, it will lead us to an interesting topological perspective on the double copy, that we will see in the following section.

3.2.1 The gauge field of a non-abelian monopole

Here we construct the most general gauge field describing a singular magnetic monopole in non-abelian gauge theory with gauge group G. The procedure follows that found in ref. [92]. We focus on static solutions with magnetic charge but no electric charge, such that $F_{0i} = 0$ and we may pick a gauge in which $A_0 = 0$. Furthermore, let us assume that we can expand the remaining Cartesian components of the gauge field in inverse powers of the radial coordinate r, such that

$$A_i = \frac{a_i(\theta, \phi)}{r} + \mathcal{O}(r^{-2}). \tag{3.47}$$

Here $A_i = A_i^a T^a$ and $a_i = a_i^a T^a$ are valued in the Lie algebra of G. This generates a magnetic field

$$B_i = -\frac{1}{2}\epsilon^{ijk} F_{jk}, \tag{3.48}$$

where F_{ij} are the spatial components of the Yang-Mills field strength. For a point-like monopole solution we want a magnetic field with r^{-2} dependence, corresponding to a Coulomb-like charge. As the magnetic field is given in terms of the field strength, which contains partial derivatives of the gauge field and a non-linear term, the highest-order terms in the gauge field that will contribute go like r^{-1}. Thus we can omit the $\mathcal{O}(r^{-2})$ terms in eq. (3.47).

To avoid any issues arising from the unavoidable singularity at the origin, we consider the field only for $r > r_0$, where r_0 is the small but non-vanishing radius of a sphere centred on the monopole. To simplify the analysis, we now begin by choosing a gauge in which $A_r = 0$. This can be done by finding a non-abelian gauge transformation such that

$$A_r \to U A_r U^{-1} + U \partial_r U^{-1} = 0. \tag{3.49}$$

A solution to this expression is

$$U^{-1}(r, \theta, \phi) = \mathcal{P} \exp\left[\frac{i}{g}\int_{r_0}^{r} dr' A_r(r', \theta, \phi)\right], \tag{3.50}$$

where $\mathcal{P}$ denotes path ordering [92]. This describes an integration along radial lines, with the lower bound protecting the integral from the singularity at the origin. A similar procedure can be implemented for A_θ by integrating along lines of constant r and ϕ, resulting in a gauge in which $A_\theta = 0$. This leaves A_ϕ as the only non-zero component of the potential, and we may determine its form from the equations of motion. As stated previously, the monopoles considered here are time-independent with no electric charge, so the F_{0i} components of the field strength vanish. Furthermore, the assumptions made thus far imply that at large distances A_ϕ is independent of r. The only non-vanishing component of the field strength is therefore

$$F_{\theta\phi} = \partial_\theta A_\phi. \tag{3.51}$$

The Yang-Mills equations of motion in spherical coordinates are

$$\partial_\mu \sqrt{\eta} F^{\mu\nu} - ig[A_\mu, \sqrt{\eta} F^{\mu\nu}] = 0, \tag{3.52}$$

where η is the absolute value of the determinant of the Minkowski metric. The field equations give rise to two non-trivial equations of motion:

$$\partial_\theta \sqrt{\eta} F^{\theta\phi} = 0, \tag{3.53}$$
$$\partial_\phi \sqrt{\eta} F^{\phi\theta} - ig[A_\phi, \sqrt{\eta} F^{\phi\theta}] = 0. \tag{3.54}$$

The first equation enforces

$$\partial_\theta \left(\frac{1}{\sin\theta} \partial_\theta A_\phi \right) = 0, \tag{3.55}$$

which admits a general solution

$$A_\phi = M(\phi) + \frac{Q_M(\phi)}{4\pi} \cos\theta. \tag{3.56}$$

Here $M(\phi)$ and $Q_M(\phi)$ are Lie algebra valued matrices, and the factor of $1/4\pi$ has been included by convention. As in the abelian case, it is not possible for this potential to be well-defined for all θ and we will once again end up with a Dirac string. Choosing this to lie along the negative z-axis, we require that A_ϕ vanishes at $\theta = 0$ to avoid another singularity at the north pole. This fixes $M(\phi)$ to be

$$M(\phi) = -\frac{Q_M(\phi)}{4\pi}. \tag{3.57}$$

Now using eqs. (3.56) and (3.57) in the second equation of motion, eq. (3.54), yields

$$\partial_\phi Q_M(\phi) = 0, \tag{3.58}$$

and hence Q_M is a constant matrix. The general solution for the gauge field is therefore

$$A^{\mathrm{N}}_\phi = \frac{Q_M}{4\pi}(\cos\theta - 1), \tag{3.59}$$

where the label N denotes that this is non-singular everywhere in the northern hemisphere. From eq. (3.48), we find that this generates a magnetic field

$$B_i = \frac{Q_M}{4\pi}\frac{\hat{r}_i}{r^2}. \tag{3.60}$$

This is precisely the form of eq. (3.8). It differs in the fact that the charge Q_M is now a matrix in the Lie algebra of G, rather than a scalar. This will be referred to as the magnetic charge matrix.

The form of the gauge field in eq. (3.59) has important consequences for the double copy. We see that it is just a Dirac monopole (eq. (3.12)) dressed with a non-abelian charge matrix. This is exactly the form of eq. (3.45), which is already known to be the single copy of Taub-NUT spacetime. We may therefore conclude that it is *always* possible to choose a gauge in which the double copy of the non-abelian monopole is straightforward, regardless of the gauge group within which it is embedded. Thus, for arbitrary gauge groups, the double copy of a non-abelian monopole is the pure NUT solution in gravity.

3.2.2 Classifying the monopole solutions

The monopole gauge field in eq. (3.59) is defined with the Dirac string aligned along the negative z-axis. As in the abelian case the location of this string singularity is arbitrary, and can be moved to align with any curve via an appropriate gauge transformation. To define a non-abelian gauge field that is non-singular everywhere away from the origin, we can follow the process prescribed by the Wu-Yang construction for abelian monopoles. That is, we introduce a second gauge field for which the Dirac string is located in the opposite direction, aligning with the positive z-axis,

$$A^{\mathrm{S}}_\phi = \frac{Q_M}{4\pi}(\cos\theta + 1). \tag{3.61}$$

This is therefore non-singular everywhere in the southern hemisphere. The northern and southern fields are related by a non-abelian gauge transformation

$$A^{\mathrm{N}}_\mu = S(\phi)A^{\mathrm{S}}_\mu S(\phi) - \frac{i}{g}S(\phi)\partial_\mu S^{-1}(\phi), \tag{3.62}$$

where

$$S(\phi) = \exp\left[\frac{igQ_M}{2\pi}\phi\right]. \tag{3.63}$$

Requiring this to be single valued, such that $S(0) = S(2\pi)$, leads to

$$e^{igQ_M} = I, \tag{3.64}$$

where I is the identity element in G. This is the non-abelian generalisation of the Dirac quantisation condition in eq. (3.22).

Let us now study the form of the magnetic charge matrix Q_M. Recall that the Cartan subalgebra of G is the largest subset of mutually commuting generators [101]. Without loss of generality, we can always write Q_M as a linear combination of the generators H_i of the Cartan subalgebra [102], such that

$$Q_M = 4\pi \sum_{i=1}^{r} w_i^* H_i, \tag{3.65}$$

where r denotes the rank of G, and the coefficients w_i^* are known as *magnetic weights*. The reason for the star notation will become apparent soon. Let us collect the magnetic weights into a *magnetic weight vector* $\boldsymbol{w}^*$. In a basis where Cartan subalgebra generators H_i are simultaneously diagonalisable, the diagonal elements of Q_M will take the form $4\pi\boldsymbol{w}^* \cdot \boldsymbol{w}$. The vector $\boldsymbol{w}$ is a *weight vector* in the representation in which the Cartan generators are expressed. Taking this form of the charge matrix in the generalised quantisation condition, eq. (3.64), reduces it to a condition on the weights and magnetic weights:

$$\boldsymbol{w}^* \cdot \boldsymbol{w} = \frac{n}{2g}, \quad n \in \mathbb{Z}. \tag{3.66}$$

The magnetic weights $\boldsymbol{w}^*$ are the weights of a Lie group G^* that is dual to G. To see this, consider the well-known fact that for a given representation the roots $\boldsymbol{\alpha}$ and weights $\boldsymbol{w}$ must always satisfy [101]

$$\frac{2\boldsymbol{w} \cdot \boldsymbol{\alpha}}{\boldsymbol{\alpha}^2} = N, \quad N \in \mathbb{Z}. \tag{3.67}$$

Hence, a solution to eq. (3.66) is

$$\boldsymbol{w}^* = \sum_i n_i \boldsymbol{\alpha}^{(i)*} = \sum_i n_i \frac{\boldsymbol{\alpha}^{(i)}}{|\boldsymbol{\alpha}^{(i)}|^2}, \tag{3.68}$$

where n_i are integers and $\boldsymbol{\alpha}^* = \boldsymbol{\alpha}/|\boldsymbol{\alpha}|^2$ are the roots of the dual group G^*. We are thus left with two separate systems of root lattices. Following the terminology of ref. [92], we refer to G and G^* as electric and magnetic gauge groups respectively. The electric gauge group G has roots $\boldsymbol{\alpha}$ and weights $\boldsymbol{w}$ which correspond to representations of the fields present in the theory. The magnetic group G^* has roots $\boldsymbol{\alpha}^*$ and weights $\boldsymbol{w}^*$ which correspond to the possible magnetic charges. If both groups share the same Lie algebra, then their root vectors will differ only by a rescaling. Furthermore, if G is the universal covering of the algebra then all possible magnetic weights are specified by eq. (3.68). In general more solutions will exist. A number of examples of electric gauge groups G

and their magnetic duals G^* can be found in ref. [102].

Non-abelian monopole solutions are characterised by the magnetic charge matrix, and we have now seen that the possible values of this matrix are fixed by the weights of the magnetic gauge group G^*. Naively, it would appear that there should be an infinite number of possible magnetic charges, corresponding to arbitrary weight vectors in the dual root lattice. However, it turns out that these do not all correspond to physically distinct or stable monopoles. Now that the problem of finding magnetic charge matrices has be formulated in the language of roots and weights, known results from group theory can be used to constrain the possible solutions.

Firstly, any magnetic weight $\boldsymbol{w}^*$ necessarily gives rise to a second weight $-\boldsymbol{w}^*$. Hence, one might assume the existence of two separate monopole solutions with charge matrices Q_M and $-Q_M$. However, the orientation of the generators within the Cartan space can be transformed such that $H_i \to -H_i$ under a global gauge transformation. The monopoles defined by charge matrices differing by a sign can therefore be identified as physically equivalent. This identification can be generalised to include other symmetries of the root and weight system. Consider two weights $\boldsymbol{r}$ and $\boldsymbol{s}$ such that $\boldsymbol{r} \neq \boldsymbol{s}$. We may always define a third weight via a Weyl reflection of $\boldsymbol{s}$ with respect to $\boldsymbol{r}$ [103],

$$\boldsymbol{w_r} : \boldsymbol{s} \to \boldsymbol{w_r}(\boldsymbol{s}) = \boldsymbol{s} - 2\boldsymbol{r}\frac{\boldsymbol{s}\cdot\boldsymbol{r}}{|\boldsymbol{r}|^2}. \tag{3.69}$$

This transformation describes the reflection of $\boldsymbol{s}$ with respect to the hyperplane in the root space that is perpendicular to $\boldsymbol{r}$, and gives rise to the root $\boldsymbol{w_r}(\boldsymbol{s})$. Weyl reflections of magnetic weights transform the magnetic charge matrix in an identical way to gauge transformations of the Cartan generators. Hence, two magnetic charge matrices defined by magnetic weights related by Weyl transformations give rise to physically equivalent monopoles.

Further restrictions arise from the structure of the lattice generated by the magnetic weights. Weight lattices may be divided into sublattices, where each sublattice contains weights that differ by an integral sum of roots. For two weights in separate sublattices this cannot be the case. It can be shown that there is a one-to-one correspondence between the magnetic weight sublattices and the elements of the first homotopy group of G [92]. Only monopole solutions with magnetic weights existing in separate sublattices can be considered to be physically distinct. Furthermore, stability analysis carried out in refs. [89, 91] found that within each sublattice the only stable solution is that with the minimum value of $\operatorname{tr} Q_M^2$. All other solutions corresponding to magnetic weight vectors within a given sublattice will decay to the solution corresponding to this value. Thus, all monopoles that exist within the same sublattice as the origin, corresponding to the zero-charge configuration, are topologically equivalent to the vacuum. After factoring out these considerations, we find that the number of physically distinct stable monopoles is equal to the number of elements of the first homotopy group of G, with one subtracted corresponding to the zero-charge vacuum solution.

It is important to note that this instability of monopoles with non-minimal magnetic weights within a given sublattice necessitates a careful specification of the gauge groups. If either the electric or magnetic gauge group is the universal covering group $\tilde{G}$ of the Lie algebra, then the other will be the adjoint group $\tilde{G}/K$, where K is the centre of $\tilde{G}$. Thus if we consider Yang-Mills with $G = \mathrm{SU}(N)$, the magnetic group will be $G^* = \mathrm{SU}(N)/\mathbb{Z}_N$. However, all weights of $\mathrm{SU}(N)/\mathbb{Z}_N$ lie in the same sublattice as the weight at the origin of the lattice. All monopole solutions for $G = \mathrm{SU}(N)$ are therefore dynamically unstable and will reduce to the vacuum solution. This, however, is not the case for $G = \mathrm{SU}(N)/\mathbb{Z}_N$ and $G^* = \mathrm{SU}(N)$, as the N sublattices of the weight lattice of $\mathrm{SU}(N)$ allow for $N-1$ stable monopole solutions. This structure is closely related to the topological properties of monopole solutions, and we turn to this now.

3.3 A topological view of the double copy

3.3.1 The topology of abelian monopoles

In defining a gauge field for an abelian monopole that is non-singular everywhere away from the origin we required the introduction of two separate fields that are related via a gauge transformation in the region in which they overlap, which we take to be the equator. The patching together of two fields on the equator is in fact an inherently topological construction, as can be made apparent via a fibre bundle interpretation of the monopole gauge fields.

Yang-Mills theory admits an elegant geometrical formulation in terms of principal fibre bundles. The gauge field is interpreted as a connection on a manifold that looks locally like the product space $\mathcal{M} \times G$. Here $\mathcal{M}$ is the base space, which for Yang-Mills on a flat background is simply Minkowski spacetime. G is the fibre, which in the case of principal fibre bundles aligns with the gauge group. While the fibre bundle appears locally to be the product space $\mathcal{M} \times G$, its global structure may differ. Thus, for a given gauge group, more than one fibre bundle may possess the same local structure. This description reduces the classification of topologically non-trivial gauge fields to that of distinct fibre bundles.

In the case of the Dirac monopole, that is necessarily singular at the origin, the base space is Minkowski spacetime with the origin removed. The spatial part of this manifold is thus $\mathbb{R}^3 - \{0\}$, which is homeomorphic to a sphere S^2. Note that it is a well known result that it is not possible to define a vector field on S^2 without it being singular at a point. This itself explains the presence of the Dirac string singularities. As we are here considering abelian monopoles, the fibre is simply $\mathrm{U}(1) = S^1$. Restricting ourselves to time independent monopoles, our principal bundle is thus locally the product space $S^2 \times S^1$.

We now define two charts $\{L_\mathrm{N}, L_\mathrm{S}\}$ that provide an open covering of the base space S^2,

$$L_\mathrm{N} = \{(\theta, \phi) \mid \theta \in [0, \pi/2],\ \phi \in [0, 2\pi)\} \tag{3.70}$$

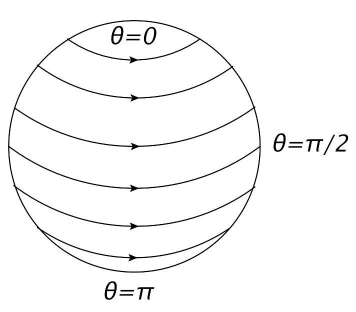
$$L_\mathrm{S} = \{(\theta, \phi) \mid \theta \in [\pi/2, \pi],\ \phi \in [0, 2\pi)\} \tag{3.71}$$

These correspond to the northern and southern hemispheres of the base space respectively. The gauge fields of eq. (3.12) provide local connections on their respective hemispheres. On the equator, where the two fields overlap, they are related by a gauge transformation. Consider the family of circles $C(\theta)$ on the two-sphere at infinity, as shown in figure 3.2(a). These circles are characterised by the polar angle θ such that the entire sphere is covered as θ goes from 0 to π. To each loop one may associate a U(1) group element:

$$U(\theta) = \exp\left(ig \oint_{C(\theta)} dx^\mu A_\mu \right). \tag{3.72}$$

This is an example of a Wilson loop, a fact that we will return to in the following chapter. The group element $U(\theta)$ defines a map from a circle $C(\theta)$ on the base space to the fibre, $U : S^1 \to S^1$.

Let us consider the behaviour of this group element as θ varies from 0 to π. At $\theta = 0$, $C(0)$ is an infinitesimally small loop centred on the north pole, and thus corresponds to the identity element of the gauge group. In the interval $\theta \in [0, \pi/2]$, $U(\theta)$ then traces out a smooth curve in the gauge group. However at $\theta = \pi/2$, corresponding to the equator, there is a discontinuity where the gauge field instantaneously switches from its northern form to its southern form. Then for $\theta \in [\pi/2, \pi]$, $U(\theta)$ continues in a smooth fashion until $\theta = \pi$, where $U(\pi)$ is again the identity element corresponding to an infinitesimally small loop now centred on the south pole. Let A correspond to the point defined by $U(\pi/2)$ in terms of A^N_μ, and B correspond to $U(\pi/2)$ in terms of A^S_μ. Physically, eq. (3.72) represents the Aharonov-Bohm phase factor picked up by the wavefunction of an electrically charged particle upon moving around the closed loops $C(\theta)$ [104]. Requiring that this is single valued at all points on the equator corresponds



(a) (b)

Figure 3.2: (a) A family of curves parameterised by θ. The north and south poles correspond to $\theta = 0$ and $\theta = \pi$ respectively, such that the entire sphere is covered as θ varies from 0 to π. (b) The closed loop in the gauge group associated with the curves in (a) via eq. (3.72). The discontinuity between A and B corresponds to the gauge field changing from its northern to its southern form at the equator. However, for correct patching of the gauge fields, A and B must be associated with equivalent group elements.

to the requirement that the points A and B correspond to equivalent group elements. Thus, as $C(\theta)$ covers the entire sphere for $\theta \in [0, \pi]$, $U(\theta)$ traces out a *closed* curve in the gauge group. This is depicted in figure 3.2(b).

From this discussion we see how the presence of the monopole within the sphere gives rise to a topological obstruction, with the magnetic charge classified by the possible ways of joining the points A and B in the gauge group G. This corresponds to the set of equivalence classes of topologically inequivalent closed loops in G, which is just the first homotopy group $\pi_1(G)$. In the present case we have

$$\pi_1(\mathrm{U}(1)) = \mathbb{Z}. \tag{3.73}$$

This is the simple result that any closed loop can wrap around a circle an integer number of times. There are thus an infinite number of different monopole solutions corresponding to the discretely different U(1) principal bundles. By imposing that the points A and B must correspond to equivalent group elements, we require

$$U_{\mathrm{N}}\big|_{\theta=\frac{\pi}{2}} = e^{2\pi i n} U_{\mathrm{S}}\big|_{\theta=\frac{\pi}{2}}, \qquad n \in \mathbb{Z} \tag{3.74}$$

where U_{N} and U_{S} correspond to eq. (3.72) written in terms of the northern and southern hemisphere gauge fields respectively. As eq. (3.72) is a U(1) element, this is simply the statement that

$$e^{i\phi} = e^{i(\phi + 2\pi n)}. \tag{3.75}$$

Inserting the gauge fields of eqs. (3.12), eq. (3.74) translates to

$$ig \oint_{C(\frac{\pi}{2})} dx^\mu \left(A_\mu^{\mathrm{N}} - A_\mu^{\mathrm{S}}\right) = 2\pi i n, \tag{3.76}$$

which, upon performing the integrals, enforces the Dirac quantisation condition of eq. (3.22). While this construction is more obscure for the case of non-abelian monopoles, the general philosophy carries over. We turn to this situation next.

3.3.2 The topology of non-abelian monopoles

Let us now generalise the preceding discussion to the topology of non-abelian monopoles. The initial set-up carries over from the abelian case. We remove the singularity at the origin such that spatial slices of the spacetime surrounding the monopole are topologically equivalent to S^2. This requires at least two coordinate patches, which we take to be the northern and southern hemispheres. Once again we consider the set of curves $C(\theta)$ shown in figure 3.2(a), which are now associated with elements of the non-abelian gauge group G via

$$U(\theta) = \mathcal{P} \exp\left(ig \oint_{C(\theta)} dx^\mu A_\mu\right), \tag{3.77}$$

where $\mathcal{P}$ denotes path ordering of the gauge fields along the curve, which is required as we are integrating over a non-abelian gauge field. Equation (3.77) provides a map from the the curves $C(\theta)$ to the gauge group, such that

$$U : S^1 \to G. \tag{3.78}$$

As θ varies from 0 to π this traces out a curve in the gauge group, with $\theta = 0, \pi$ corresponding to the identity element. At $\theta = \pi/2$ there is a discontinuity generated by the gauge field switching from its northern to southern form, where these fields are given in eqs. (3.59) and (3.61) respectively. Due to the clear similarities with the abelian case, this situation may also be represented by figure 3.2(b), where A and B represent the points corresponding to $U(\pi/2)$ in terms of the northern and southern field respectively. We require that A and B correspond to equivalent group elements. Let H be the group of transformations between A and B such that they are identified as equivalent, then the patching condition for non-abelian fields can be written as

$$\mathcal{P} \exp\left(ig \oint_C dx^\mu A_\mu^{\mathrm{N}}\right) = U_H \left[\mathcal{P} \exp\left(ig \oint_C dx^\mu A_\mu^{\mathrm{S}}\right)\right], \tag{3.79}$$

where U_H is an element of H and $C \equiv C(\pi/2)$ is the curve coinciding with the equator. As H is simply the group of transformations between the points A and B such that they are equivalent group elements, each correct patching in eq. (3.79) corresponds to a different closed loop within the gauge group, the set of which forms the first homotopy group $\pi_1(G)$. We refer to eq. (3.79) as the *patching condition* for non-abelian gauge fields. The patching condition encodes the non-trivial topology of the non-abelian monopole. For a gauge group $G = \tilde{G}/K$, when $\tilde{G}$ is the universal coving group, the relevant homotopy group is

$$\pi_1(\tilde{G}/K) = K. \tag{3.80}$$

For pure gauge theory we have $G = \mathrm{SU}(N)/\mathbb{Z}_N$ and thus $\pi_1(G) = \mathbb{Z}_N$. Here the monopole corresponding to the identity element is topologically equivalent to the vacuum solution, and is thus unstable. We therefore find $N-1$ stable non-abelian monopole solutions, consistent with the discussion of the previous section.

Let us consider the simple case of $G = \mathrm{SU}(2)/\mathbb{Z}_2 = \mathrm{SO}(3)$. In this case we have

$$\pi_1(\mathrm{SU}(2)/\mathbb{Z}_2) = \mathbb{Z}_2, \tag{3.81}$$

such that there are two distinct monopole configurations. The trivial patching is topologically equivalent to the vacuum solution. We are thus left with a single stable monopole configuration, corresponding to the patching condition

$$\mathcal{P}\exp\left(ig\oint_C dx^\mu A^{\mathrm{N}}_\mu\right) = -I\left[\mathcal{P}\exp\left(ig\oint_C dx^\mu A^{\mathrm{S}}_\mu\right)\right], \tag{3.82}$$

where I is the identity element in the gauge group. For the case of an abelian gauge group, the condition of eq. (3.79) reduces to an abelian patching condition. This can be easily seen by noting that a general element of H is

$$U_H = e^{2\pi i n}, \quad n \in \mathbb{Z}, \tag{3.83}$$

corresponding to the fact that $\pi_1(\mathrm{U}(1)) = \mathbb{Z}$. Taking this in eq. (3.79) and noting that the path ordering is no longer necessary for abelian fields, we reacquire eq. (3.74).

The form of the non-abelian patching condition in eq. (3.79) has interesting consequences for the double copy. In the abelian case, the patching condition is equivalent, up to a constant factor, to the first Chern number

$$c_1 = \frac{1}{4\pi}\int_\Sigma F_{\mu\nu} d\Sigma^{\mu\nu}, \tag{3.84}$$

where Σ is the closed surface, in our case S^2. This is to be expected as the first

Chern number classifies the topology of U(1) principal bundles. However, for other gauge groups, this relation will not necessarily carry over and a different topological invariant will need to be found. For example, in the case of SU(N) monopoles, the first Chern number vanishes as the generators are traceless. This creates an issue in trying to identify global characteristics which match up on either side of the double copy correspondence. We have seen that monopoles in arbitrary gauge groups double copy to the same gravity solution, and thus it appears that the double copy is in some sense blind to the gauge group. On the other hand, this is clearly not the case for topological invariants which are fully dependent on the gauge group. We see here, however, that it is not actually necessary to identify a topological invariant, as the patching condition of eq. (3.79) completely specifies the non-trivial topology of the monopole solutions, and its form remains constant regardless of the gauge group within which the monopole is embedded. Furthermore, we will now see that eq. (3.79) has a well-defined gravitational counterpart.

3.3.3 The topology of Taub-NUT spacetime

To develop a topological description of the Taub-NUT solution that relates to the preceding discussion in gauge theory, it will be useful to obtain the periodicity condition in eqs. (3.32) and (3.34) via an alternative approach. This involves considering the time shift experienced by a test particle upon moving around a closed loop, known as the time holonomy. This is defined by [105]

$$|\Delta t| = \oint_C dx^i \, \frac{g_{0i}}{g_{00}}, \tag{3.85}$$

where C is a closed loop and $i \in \{1, 2, 3\}$ denotes spatial indices. Recall that in section 3.1.2 we introduced a northern and southern form for the pure NUT metric in eq. (3.31), in analogy to the Wu-Yang construction of the Dirac monopole. Taking the $r \to \infty$ limit in these metrics leaves them in the form

$$g_{\mu\nu}^{N,S} = \eta_{\mu\nu} + \kappa h_{\mu\nu}^{N,S}, \tag{3.86}$$

where the only non-zero components of the graviton fields are

$$h_{0\phi}^{N,S} = 2N\left(\cos\theta \mp 1\right), \quad h_{\phi\phi}^{N,S} = 4\kappa N^2 \left(\cos\theta \mp 1\right)^2. \tag{3.87}$$

Thus, the time holonomy reduces to a form completely in terms of the graviton fields,

$$|\Delta t^{N,S}| = \kappa \oint_C dx^i h_{0i}^{N,S}. \tag{3.88}$$

Consider evaluating the time holonomy around one of the curves of constant θ shown in figure 3.2. Due to the ϕ independence of $h^{N,S}_{0\phi}$ this takes the simple form

$$|\Delta t^{N,S}(\theta)| = \kappa \oint_{C(\theta)} d\phi\, h^{N,S}_{0\phi} d\phi = 2\pi\kappa h^{N,S}_{0\phi}. \tag{3.89}$$

The time shifts form a subgroup of the general group of diffeomorphisms that act in general relativity. At the equator, where the northern and southern fields overlap, we require that the time holonomies arising from each field coincide. However, evaluating the difference on the equator we find that

$$|\Delta t^{S}(\pi/2)| - |\Delta t^{N}(\pi/2)| = 8\pi\kappa N. \tag{3.90}$$

Thus the periodicity condition given in eqs. (3.32) and (3.34) is required for the appropriate patching of the fields around the equator.

There is an similar construction which makes contact with the original argument given by Dirac for the quantisation of electric charge in the presence of an abelian monopole [106]. In ref. [107] analogies between linearised gravity and magnetic monopoles were studied. This was done by considered the phase picked up by the wavefunction of a non-relativistic test particle of mass m as it traverses a closed loop C in a non-trivial gravitational background. It was found that if $h_{00} = 0$, then in the weak-field limit this phase is given by

$$\Phi = \exp\left[i\kappa m \oint_C dx^i h_{0i}\right]. \tag{3.91}$$

Note that this is simply the exponentiated time holonomy of eq. (3.88)

$$\Phi = e^{im|\Delta t|}. \tag{3.92}$$

For this to be well-defined on the equator, where the northern and southern fields overlap, the difference in the phases evaluated with each field must be an integer multiple of 2π, such that

$$\kappa m \oint_C dx^i \left[h^N_{0i} - h^S_{0i}\right] = 2\pi n, \quad n \in \mathbb{Z}. \tag{3.93}$$

By substituting the explicit forms for the pure NUT metrics in eq. (3.87), this yields

$$m\kappa N = \frac{n}{4}, \tag{3.94}$$

implying that the mass of the test particle is quantised. Reference [107] was unsure about how to interpret this mass. However, the periodicity of the time coordinate makes this clear. The mass m describes the energy of a static wavefunction in the presence of the NUT charge. If the time coordinate is compact with period t_0, this implies that the mass m is quantised according to

$$\Delta m = \frac{2\pi}{t_0}. \tag{3.95}$$

Equation (3.94) implies

$$\Delta m = \frac{1}{4\kappa N_0}, \tag{3.96}$$

so that combining this with eq. (3.95) yields the period of the time coordinates as required.

The periodic time coordinate gives the Taub-NUT solution the topology of a $U(1)$-bundle over S^2. On the equator we require that the northern and southern metrics are correctly patched together. Consider forming the following quantity from the time holonomies of eq (3.88)

$$\Phi(\theta) = \exp\left(i\kappa \oint_{C(\theta)} dx^\mu h_{0\mu}\right), \tag{3.97}$$

where $C(\theta)$ are again the family of curves covering the two-sphere at infinity in figure 3.2(a), and we have used the fact that $h_{00} = 0$. Due to the periodicity of the time coordinate, this provides a map from the base spacetime into the fibre, $\Phi : S^1 \to S^1$. We can now follow the same process as in the monopole case. At $\theta = 0$, $C(0)$ describes an infinitesimally small loop centred on the north pole, and thus eq. (3.88) corresponds to the identity element of the group. As θ varies from 0 to π a curve is traced out in the $U(1)$ group manifold, with $C(\pi)$ once again corresponding to the identity element. There is, however, a discontinuity at $\theta = \pi/2$ where the graviton field switches from its northern to its southern form. This can once again be represented pictorially by figure 3.2(b). We therefore must patch the graviton fields at the equator such that

$$\exp\left(i\kappa \oint_C dx^\mu h^N_{0\mu}\right) = U_H \left[\exp\left(i\kappa \oint_C dx^\mu h^S_{0\mu}\right)\right], \tag{3.98}$$

where U_H is an element of the group of transformations between points A and B in figure 3.2(b), and $C \equiv C(\pi/2)$. The elements of the group H represent equivalence classes of closed loops in the group manifold, and are therefore in one-to-one correspondence with the elements of the first homotopy group. As the group here is $U(1)$, an example of a general element U_H can be found in eq. (3.83). Eq. (3.98) is therefore the patching condition for Taub-NUT spacetime, which classifies the non-trivial topology of the solution.

3.3.4 Discussion

Here we have developed a description of the topology of monopoles, both abelian and non-abelian, and Taub-NUT spacetime in terms of patching conditions. The patching condition for the Taub-NUT solution in eq. (3.98) is a direct analogue of the patching condition for the non-abelian monopole in eq. (3.79). As these conditions encode the non-trivial topology of their respective objects, and these objects are indeed dual under the double copy, they represent global information that is preserved under the double copy correspondence. Thus, eqs. (3.79) and (3.98) act as an example of a global rather than local statement of the double copy. It is not surprising that the Taub-NUT condition takes an abelian-like form, while the monopole case can be abelian or non-abelian. This precisely mimics the local statement of the double copy in which the colour structure, within which the non-abelian nature of the monopole is encapsulated, is removed upon taking the double copy. Hence, as in the local case, an abelian-like construction is always obtained in the gravity theory, regardless of the gauge group one starts with in the gauge theory.

A curious property of the patching conditions of eqs. (3.79, 3.98) is that they are phrased in terms of certain Wilson line operators. One might then ask whether the general form of such operators can be interpreted from the perspective of the double copy. This indeed turns out to be the case, as we will see in the next chapter. However, in answering this question we will be left with another, associated with the closely related concept of holonomy.

Chapter 4

Wilson lines, holonomy, and the double copy

In the previous chapter we developed an example of the classical double copy in which both the global and local structure of certain solutions was identified under the correspondence. This topological perspective represents one possible route into the non-perturbative structure of the double copy. In this chapter we will take an alternative albeit closely related path, centred on the geometrical concept of holonomy.

Loosely speaking, holonomy describes the transformation of certain mathematical objects after parallel transport in a non-trivial background. The set of all such transformations forms a group structure, known as the holonomy group. In both gauge theory and gravity, the geometrical origins of the holonomy are closely analogous. Both pure Yang-Mills and pure Einstein gravity can be formulated in the language of fibre bundles, where the fibres correspond to the gauge group and tangent space respectively. In both cases, parallel transport in the base space induces transport in the fibre. Holonomy in gauge theory and gravity can thus be traced back to a common geometrical origin, and one might therefore hope that there exists some explicit double copy between them. Indeed this is a tantalising prospect, given that the geometrical foundations of the exact classical double copy remain obscure. We will see in the following that this turns out not to be the case, and that this fact is closely related to a double copy interpretation for Wilson lines in gauge theory and gravity.

Mathematically, holonomy is quantified by the integration of a connection over a closed curve in either gauge theory or gravity. In the case of gauge theory, such an operator is a specific case of what is more generally referred to as a Wilson line, for which the integration contour can be arbitrary. In gravity the situation is not so clear cut, as historically two operators have been taken as the definition of a gravitational Wilson line. The first relies on the geometrical origins outlined in the previous paragraph. By relaxing the integration contour of the holonomy in a gravitational theory, such that one has the integral of e.g. the Christoffel connection over an arbitrary curve, one obtains an operator that appears to be the sensible gravitational analogue of a Wilson line in gauge theory. Indeed, such an operator was considered in e.g. refs. [108–113]. However, in refs. [109, 110, 112], the perturbative behaviour of this operator was found to be in stark contrast to its gauge theory counterpart. The second operator labelled as the gravitational Wilson line involves an integration of the metric itself, as considered in e.g. refs. [112, 114–117]. This operator can be interpreted as the phase picked up by the wavefunction of a test particle as it traverses the integration contour, and is in this sense a *physical* analogue of the gauge theory Wilson line which arises in the description of the Aharanov-Bohm effect. Furthermore, it plays an analogous role to the gauge theory Wilson line in the study of amplitudes [116, 118, 119], where it gives rise to results that can be independently obtained via colour-kinematics duality and the double copy [1, 120–122]. From this discussion it appears that the operator that is considered to be analogous to the gauge theory Wilson line is dependent on the perspective one takes. From a geometrical perspective it is the integration over the connection, which gives rise to the gravitational holonomy, whereas from a physical perspective it is the integration over the metric.

Here we will develop a notion of the double copy for all of the operators mentioned above, thereby furnishing this fairly confusing cast of characters with some form of organisational structure. In doing so we will be forced to introduce yet another operator, although this will in turn allow us to investigate some interesting non-perturbative aspects of the double copy. We will begin by reviewing some basic facts about Wilson lines and holonomy in gauge theory and gravity, which we hope will help clarify the distinctions between the many operators mentioned here. We will then go on to develop a notion of the double copy for Wilson line operators themselves. We will justify this by the relation to both the infrared structure of gauge theory and gravity and the topological discussion of the previous chapter. This will furthermore put the results of the previous chapter on firmer ground. As a result of this discussion, the gravitational holonomy will be left without a double copy partner. This will be corrected in the following section, in which we will identify the single copy of the gravitational holonomy and justify its form via relations to amplitudes and exact classical solutions. Finally, we will study the explicit form of the gravitational holonomy and its single copy, along with their resultant groups, for a number of solutions that are known to be related by the double copy.

4.1 Wilson lines, Wilson loops, and holonomy

4.1.1 Riemannian holonomy

Consider an m-dimensional (pseudo-)Riemannian manifold M, and the set of all closed loops at a point p on M. Given a connection on the manifold, we may parallel transport a vector defined in the tangent space at p, $X \in T_pM$, around one of the loops. The resulting vector, $X' \in T_pM$, will not in general be equal to the original vector. Thus, the loop and the connection define a linear transformation at the point p, the set of which form the holonomy group at p, $H(p)$.

To study the elements of the holonomy group we can introduce the parallel propagator. If $X \in T_pM$ is parallel transported along a curve $\gamma : \lambda \to x^\mu(\lambda)$, then the components of X at any given point on the curve can be related to those at p via a matrix ${\Phi^\mu}_\nu(x, x_0)$, where x_0 is the coordinate at p:

$$X^\mu(x) = {\Phi^\mu}_\nu(x, x_0) X^\nu(x_0). \tag{4.1}$$

The transformation matrix, referred to as the *parallel propagator*, is found by solving the parallel transport equation,

$$\frac{d}{d\lambda} X^\mu + \frac{dx^\nu}{d\lambda} \Gamma^\mu_{\nu\sigma} X^\sigma = 0, \tag{4.2}$$

which yields

$${\Phi^\mu}_\nu(x, x_0) = \mathcal{P} \exp\left[- \int_{x_0}^{x} dx^\sigma \Gamma^\mu_{\sigma\nu} \right], \tag{4.3}$$

where $\mathcal{P}$ denotes path ordering and $\Gamma^\mu_{\sigma\nu}$ is the Christoffel symbol. The path ordering indicates that in expanding the exponential the Christoffel symbols are to be ordered in terms of increasing values of the parameter along the curve. It is important to note that this operator is path dependent.

To see how the parallel propagator transforms under a coordinate transformation $x \to y(x)$, we write eq. (4.1) in terms of the new coordinate y,

$$X^\mu(y) = {\Phi^\mu}_\nu(y, y_0) X^\nu(y_0). \tag{4.4}$$

We then transform back to x using the transformation law for the vector components:

$$X^\mu(y) = X^\sigma(x) \left[\frac{\partial y^\mu}{\partial x^\sigma} \right]_x = {\Phi^\sigma}_\gamma(x, x_0) X^\gamma(x_0) \left[\frac{\partial y^\mu}{\partial x^\sigma} \right]_x , \tag{4.5}$$

such that eq. (4.4) becomes

$$\Phi^{\sigma}{}_{\gamma}(x,x_0)X^{\gamma}(x_0)\left[\frac{\partial y^{\mu}}{\partial x^{\sigma}}\right]_x = \Phi^{\mu}{}_{\nu}(y,y_0)X^{\rho}(x_0)\left[\frac{\partial y^{\nu}}{\partial x^{\rho}}\right]_{x_0} . \tag{4.6}$$

Here $[\alpha]_x$ denotes α evaluated at x. Now by multiplying through by $\left[\partial x^{\lambda}/\partial y^{\mu}\right]_x$ we obtain

$$\Phi^{\lambda}{}_{\rho}(x,x_0) = \Phi^{\mu}{}_{\nu}(y,y_0)\left[\frac{\partial y^{\nu}}{\partial x^{\rho}}\right]_{x_0}\left[\frac{\partial x^{\lambda}}{\partial y^{\mu}}\right]_x . \tag{4.7}$$

Finally, this can be inverted to find the transformation law for the parallel propagator:

$$\Phi^{\mu}{}_{\nu}(x,x_0) \to \Phi^{\mu}{}_{\nu}(y,y_0) = \left[\frac{\partial y^{\mu}}{\partial x^{\rho}}\right]_x \Phi^{\rho}{}_{\sigma}(x,x_0)\left[\frac{\partial x^{\sigma}}{\partial y^{\nu}}\right]_{x_0} . \tag{4.8}$$

A special case of the parallel propagator of eq. (4.3) is obtained when we consider parallel transport around a closed curve C. We then obtain the operator

$$\Phi_{\Gamma}(C) = \mathcal{P}\exp\left[-\oint_C dx^{\mu}\Gamma_{\mu}\right], \tag{4.9}$$

Note that we have written this in matrix form, with the Christoffel symbol considered as the matrix $[\Gamma_{\mu}]^{\rho}{}_{\sigma}$. This operator is referred to as the *Riemannian holonomy* or *gravitational holonomy operator*. The set of all such operators forms the *Riemannian holonomy group* $H(p)$. Holonomy can be defined in more general contexts than Riemannian manifolds, however these will be our only concern here and thus $H(p)$ will often be referred to simply as the holonomy group. Furthermore, by only considering loops that are homotopic to the identity we obtain a subgroup of $H(p)$ referred to as the *restricted holonomy group* $H_0(p)$. Naturally, if the fundamental group of the manifold is trivial, $\pi_1(M) = 0$, then the holonomy group is equal to the restricted subgroup, $H(p) = H_0(p)$.

So far we have referred to the holonomy group at a given point p on M. Consider now two points p, $q \in M$ that are connected by a curve γ, where we are now assuming that M is a connected manifold. The curve defines a map between the tangent spaces at the two points, $\tau_{\gamma} : T_pM \to T_qM$, such that

$$H(p) = \tau_{\gamma}H(q)\tau_{\gamma}^{-1}. \tag{4.10}$$

The holonomy groups at arbitrary points p and q are therefore isomorphic, allowing us to talk more generally of the holonomy group of the manifold as a whole. We denote this as $H(M)$. As a given manifold has a corresponding holonomy group, it is natural to ask whether manifolds can be classified according to their holonomy groups. This leads to Berger's classification [123].

The maximal holonomy group is $\mathrm{GL}(m,\mathbb{R})$. In general, $H(M)$ will be a subgroup of $\mathrm{GL}(m,\mathbb{R})$ and is trivial if and only if the Riemann tensor vanishes. If the connection is a metric connection and the manifold is orientable we find the following restrictions:

$$H(M) \subset \mathrm{SO}(m), \qquad \text{Riemannian manifold,} \tag{4.11}$$

$$H(M) \subset \mathrm{SO}(m-1,1), \quad \text{Lorentzian manifold.} \tag{4.12}$$

This makes sense intuitively; the metric connection preserves the length of a vector and thus the elements of the holonomy group must be orthogonal, and the orientability of the manifold ensures that their determinant is equal to one. Further reductions occur in other special cases [124], the details of which have been of much study in the mathematical literature. For a more detailed discussion of holonomy, see e.g. refs. [125, 126].

4.1.2 The spin connection holonomy

In eq. (4.2) we have defined the parallel transport equation in terms of the Christoffel connection. It will be useful for us in the following to outline an alternative description of the holonomy in terms of the *spin connection*. In general the coordinate bases of the tangent and cotangent spaces at a given point are not orthonormal. Let us instead introduce a set of orthonormal basis vectors $\{e_a\}$ at each point in spacetime, constructed from a linear combination of basis vectors of a non-orthonormal basis:

$$e_a = e_a{}^{\mu} e_{\mu}. \tag{4.13}$$

The new basis $\{e_a\}$ is referred to as the *vielbein* or *tetrad* basis, where the components $e_a{}^{\mu} \in \mathrm{GL}(m,\mathbb{R})$ are known as *vielbeins*. The inverse $e^a{}_{\mu}$ is defined such that

$$e^a{}_{\mu} e_a{}^{\nu} = \delta^{\nu}_{\mu}, \quad e^a{}_{\mu} e_b{}^{\mu} = \delta^a_b. \tag{4.14}$$

The vielbein basis vectors are orthonormal with respect to a general curved metric g, such that

$$g(e_a, e_b) = g_{\mu\nu} e_a{}^{\mu} e_b{}^{\nu} = \eta_{ab}. \tag{4.15}$$

where η is the Minkowski metric. The spacetime metric components can be written in terms of the flat metric via

$$g_{\mu\nu} = g(e_{\mu}, e_{\nu}) = g(e^a{}_{\mu} e_a, e^b{}_{\nu} e_b) = e^a{}_{\mu} e^b{}_{\nu} g(e_a, e_b) = e^a{}_{\mu} e^b{}_{\nu} \eta_{ab}, \tag{4.16}$$

where eq. (4.15) has been used in the final equality. Due to this expression it is often said that the vielbein is the square root of the metric. More importantly, it tells us that the metric is completely fixed by the vielbein. As vectors are basis independent, we may write a general vector V as

$$V = V^\mu e_\mu = V^a e_a = V^a e_a{}^\mu e_\mu, \tag{4.17}$$

from which we see that

$$V^\mu = V^a e_a{}^\mu, \quad V^a = e^a{}_\mu V^\mu. \tag{4.18}$$

Furthermore an orthonormal dual basis $\{e^a\}$ can be defined by

$$e^a = e^a{}_\mu dx^\mu, \tag{4.19}$$

such that a general curved metric takes the form

$$g = g_{\mu\nu} dx^\mu \otimes dx^\nu = \eta_{ab} e^a \otimes e^b. \tag{4.20}$$

Armed with an orthonormal basis, one can calculate the spin connection using Cartan's first structure equation. In the torsion free case, this is

$$\omega^a{}_b \wedge e^b = -de^a. \tag{4.21}$$

Inverting Cartan's structure equation then yields the spin connection in terms of the vielbein

$$(\omega_\mu)^{ab} = \frac{1}{2} e^{a\nu}\left(\partial_\mu e^b{}_\nu - \partial_\nu e^b{}_\mu\right) - \frac{1}{2} e^{b\nu}\left(\partial_\mu e^a{}_\nu - \partial_\nu e^a{}_\mu\right) - \frac{1}{2} e^{a\rho} e^{b\sigma} e^c{}_\mu \left(\partial_\rho e_{c\sigma} - \partial_\sigma e_{c\rho}\right). \tag{4.22}$$

The spin connection is both computationally and conceptually useful. Many calculations are greatly simplified through its employment, and for this reason it will be used frequently in the following sections.

Let us now formulate the holonomy in terms of the spin connection. The argument is essentially the same as that in terms of the Christoffel connection. We begin with the parallel transport equation for a vector in the vielbein basis

$$\frac{d}{dt} V^a + (\omega_\mu)^a{}_b \frac{dx^\mu}{dt} V^b = 0. \tag{4.23}$$

In direct analogy to eq. (4.3), we solve this to obtain the transformation matrix

$$[\Phi_\omega(\gamma)]^a{}_b = \mathcal{P} \exp\left[-\int_\gamma dx^\mu (\omega_\mu)^a{}_b\right]. \tag{4.24}$$

Considering a curve that connects two points p and q, this matrix acts such that

$$V_q^a = [\Phi_\omega(\gamma)]^a{}_b V_p^b. \tag{4.25}$$

Now taking the integration curve to be a closed loop C, we obtain the holonomy of the spin connection

$$\Phi_\omega(C) = \mathcal{P}\exp\left[-\oint_C dx^\mu \omega_\mu\right], \tag{4.26}$$

where we have written this in matrix notation. The spin connection holonomy is closely related to the Riemannian holonomy operator of eq. (4.9). Converting the vectors in eq. (4.1) to the vielbein basis, we find that

$$[\Phi_\omega(p,q)]^a{}_b = e^a{}_\mu(p)\,[\Phi_\Gamma(p,q)]^\mu{}_\nu\, e_b{}^\nu(q), \tag{4.27}$$

which for a close curve translates to

$$[\Phi_\omega(C)]^a{}_b = e^a{}_\mu\,[\Phi_\Gamma(C)]^\mu{}_\nu\, e_b{}^\nu, \tag{4.28}$$

where the two vielbeins on the right-hand side are evaluated at the same point. This expression has a simple physical interpretation. The Riemannian holonomy operator encodes how the components of a vector change after parallel transport around a closed loop. The spin connection does the same, but in an orthonormal basis. The two holonomy operators are therefore related by a similarity transformation, as in eq. (4.28), such that the holonomy groups defined by each operator are isomorphic. Thus, in discussing the holonomy of a manifold, one is free to use either the Riemannian or spin connection holonomies. Due to this freedom, we will often use the term gravitational holonomy to refer to either of these two operators.

For our later purposes, it will be useful to further rewrite the spin connection holonomy. The spin connection is valued in the Lie algebra of the Lorentz group, such that we may expand it in terms of Lorentz generators M^{ab} via

$$(\omega_\mu)^c{}_d = \frac{i}{2}(\omega_\mu)_{ab}(M^{ab})^c{}_d. \tag{4.29}$$

Here the normalisation arises due to the components of the generators in the spin-1 representation:

$$(M^{cd})^c{}_d = i(\eta^{ac}\delta^b_d - \eta^{bc}\delta^a_d). \tag{4.30}$$

Using the explicit form of eq. (4.29), the spin connection holonomy of eq. (4.26) is

$$\Phi_\omega(C) = \mathcal{P}\exp\left[-\oint_C dx^\mu (\omega_\mu)_{ab} M^{ab}\right]. \tag{4.31}$$

The purpose of writing the spin connection holonomy in this way will soon become clear. However, it is also nice to note in passing that this form allows for the extension of the notion of holonomy to spinors.

4.1.3 Holonomy in gauge theory

In gauge theory, the introduction of holonomy is broadly analogous to that in Riemannian geometry. The vectors we wish to transform are now fields Ψ^a identified with sections of the G-bundle, transforming in a particular representation of G, while the gauge field provides the connection. We consider a curve in the base space $\gamma : [0,1] \to M$, which passes through a point $p_0 \in M$, with coordinates x_0. The field at p_0 will be related to that at a later point $p \in M$, with coordinates x, via a rotation in the gauge space, that is induced through parallel transport along the horizontal lift of γ. That is to say, the values of the field Ψ^a at the two points along γ are related via

$$\Psi^a(x) = \Phi^a{}_b(x, x_0)\Psi^b(x_0), \tag{4.32}$$

where the matrix $\Phi^a{}_b \in G$ is obtained by solving the parallel transport equation in principle bundle, yielding

$$\Phi^a{}_b(x, x_0) = \mathcal{P} \exp\left[-g \int_{x_0}^{x} dx^\mu A_\mu\right]^a{}_b . \tag{4.33}$$

Here the gauge field $A_\mu = A^a_\mu T^a$ is valued in the Lie algebra of G, where T^a are the generators in an appropriate representation. Note that this operator is path-dependent, despite the fact that the notation does not immediately suggest this. Clearly this operator is the gauge theory analogue of the parallel propagator, and is referred to as a *Wilson line.*

If we now take the curve γ to be a loop, such that $\gamma(0) = \gamma(1) = p_0 \in M$, the Wilson line of eq. (4.33) becomes

$$\Phi_A(\gamma) = \mathcal{P} \exp\left[-g \oint_\gamma dx^\mu A_\mu\right] . \tag{4.34}$$

This operator is known as the holonomy in gauge theory. The set of all such transformations yields a subgroup of G known as the holonomy group. Note the similarity with eq. (4.9), the Riemannian holonomy operator. Both of these operators involve a path-ordered exponential involving an integral of the connection, as defined in the respective contexts. However, in the Riemannian case the operator defines a rotation in spacetime, while the gauge theory holonomy defines a rotation within the colour space associated with the gauge group.

The Wilson line operator of eq. (4.33) transforms covariantly under gauge transformations:

$$\Phi(x, x_0) \to U(x)\Phi(x, x_0)U^\dagger(x_0), \tag{4.35}$$

where $U \in G$. Note that this is analogous to the transformation of the Riemannian parallel propagator in eq. (4.8), but with the diffeomorphisms replaced with gauge transformations. It is important to emphasise the potentially confusing terminology involved here. While eq. (4.33) is a Wilson line, upon identifying the endpoints of the curve we obtain the holonomy operator, eq. (4.34), *not* a Wilson loop. A Wilson loop is obtained by taking the trace of the holonomy operator, which yields a gauge invariant quantity as a consequence of eq. (4.35).

It is clear from both the geometrical origins and the transformation properties that the parallel propagators in gauge and gravity theories, eqs. (4.33) and (4.3) respectively, are analogous to one another. For this reason the Riemannian parallel propagator is often referred to as a gravitational Wilson line. This however creates a puzzle, as there is another definition of the gravitational Wilson line that appears regularly in the literature. We turn to this now.

4.1.4 Gravitational Wilson lines

In abelian gauge theory, the Wilson line operator of eq. (4.33) has a useful physical interpretation. It represents the phase picked up by the wavefunction of a charged particle as it moves along a curve. A similar interpretation can be given to the non-abelian Wilson line, once the trace is taken to give a gauge-invariant object. Let us now consider what the analogous gravitational operator is. Here the phase can only depend on the Lorentz invariant path length of the curve, along with the mass of the test particle which is the analogue of the charge in gauge theory. We then define the *gravitational Wilson line* to be

$$\Phi_g(\gamma) = \exp\left[-im \int_\gamma d\tau \sqrt{-g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}\right]. \tag{4.36}$$

Here γ is an arbitrary curve that is parameterised by $x^\mu(\tau)$, and $\dot{x}^\mu$ represents differentiation with respect to the parameter τ. This operator has two unpleasant characteristics: the square root is cumbersome and the operator appears not to be defined for massless particles. Both of these issues can be circumvented by noticing that the exponent is simply the covariant definition of the length of a worldline parameterised by

τ. This quantity is just the action for a relativistic point particle, with the integration parameter representing the proper time. It is a well known fact that this action can be rewritten in the form [127]

$$S_{\rm pp}^{(0)} = \frac{1}{2}\int d\tau \left[\frac{1}{e(\tau)} g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu - e(\tau) m^2\right]. \tag{4.37}$$

The reason for the superscript will become apparent later in the chapter. The auxiliary field $e(\tau)$ acts as an einbein on the worldline. It is completely fixed by its equation of motion

$$\frac{\delta S}{\delta e} = -\frac{1}{2e^2} g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu - \frac{m^2}{2} = 0. \tag{4.38}$$

Solving for e and substituting the result into the action reproduces eq. (4.36) when exponentiated. However, taking $m = 0$ in eq. (4.37) now yields a perfectly well-defined action. The einbein $e(\tau)$ plays the role of a one-dimensional metric on the worldline, and thus transforms appropriately under reparameterisations. We are therefore allowed to choose a value for $e(\tau)$, as this corresponds to fixing a gauge. Taking $m = 0$ and $e(\tau) = 1$ in eq. (4.37), and exponentiating the result gives

$$\Phi_g(\gamma) = \exp\left[\frac{i}{2}\int_\gamma d\tau g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu\right]. \tag{4.39}$$

This is the analogue of the phase factor in eq. (4.36) for the case of massless particles. Let us now consider this operator in perturbation theory, in which we introduce the graviton field $h_{\mu\nu}$ via

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}. \tag{4.40}$$

To first order in κ, eq. (4.39) is then

$$\Phi_g(\gamma) = \exp\left[\frac{i\kappa}{2}\int_\gamma ds\, h_{\mu\nu}\dot{x}^\mu \dot{x}^\nu\right], \tag{4.41}$$

where we have omitted an overall multiplicative constant that will vanish in appropriately normalised vacuum expectation values of the operator.

We will primarily refer to eq. (4.41) as the gravitational Wilson line. This might seem unjustified, given that the operator of eq. (4.3), which we have called the parallel propagator, so closely mimics the geometrical content of Wilson lines in gauge theory. However, in a number of papers, eq. (4.41) has been identified as the gravitational Wilson line due to the analogous role in gravity that it plays to the gauge theory Wilson line in the description of certain physical phenomena (see e.g. refs. [112,114–116,128]). We will now briefly describe an example of one such correspondence.

The IR structure of gauge theory and gravity

Consider a general gauge theory amplitude featuring particle exchange between two of the external legs. In the soft limit, where the exchanged particle's momentum is taken to zero, the amplitude will be divergent. These IR divergences have a universal form which factors from the full amplitude, such that, schematically

$$\mathcal{A}_n = \mathcal{S} \cdot \mathcal{H}_n. \tag{4.42}$$

Here $\mathcal{A}_n$ is an n-point amplitude, $\mathcal{S}$ is known as the *soft function* and contains the IR divergences, and $\mathcal{H}_n$ is the IR-finite hard part of the interaction. The form of the soft function is known to be exponential and thus the soft corrections to the hard interaction are summed to all orders in perturbation theory (see e.g. ref. [129] for a review). We consider here the case in which the soft emission is virtual, however factorisation also occurs for real soft radiation.

For our purposes, the most interesting aspect of this story is the fact that soft functions in gauge theory can be expressed as vacuum expectation values (VEVs) of Wilson line operators, such as that in eq. (4.33), with each external leg contributing a Wilson line. Within each Wilson line, we now have an integral of the soft gauge field, where the integration contour is the physical trajectory of the hard particle from which the soft particle was emitted. Such a trajectory can be parameterised by a parameter s such that

$$x^\mu(s) = sp^\mu, \tag{4.43}$$

where p^μ is the hard momenta. For the ith external leg, the gauge theory Wilson line of eq. (4.33) then reduces to

$$\Phi_i(a, b) = \mathcal{P} \exp\left[-gp_i^\mu \int_a^b ds A_\mu\right], \tag{4.44}$$

and the soft factor is [130]

$$\mathcal{S} \sim \langle 0| \prod_i \Phi_i(0, \infty)|0\rangle. \tag{4.45}$$

This has a nice physical interpretation. The emission of a particle from an external leg would usually result in a recoil. However, as the emitted particle is soft, with vanishing momentum, no recoil occurs, and thus it is only possible for the external leg to change by a phase. This phase is exactly what the gauge theory Wilson line represents.

This construction, in which the IR structure of gauge theory factorises, can also be extended to perturbative gravity, as first shown in [131]. However, the situation in gravity is, somewhat surprisingly, far simpler than that in gauge theory. While the soft factor once again takes an exponential form, it has been shown that this exponential contains only the one-loop IR divergence [115]. Thus, the IR divergences in gravity at all orders in perturbation theory originate from the one-loop corrections. Furthermore, ref. [115] argued that the gravitational soft function could be expressed in an analogous form to that in gauge theory, eq. (4.45), as a VEV of appropriately defined operators. These operators take form

$$\Phi_i(a,b) = \exp\left[\frac{i\kappa}{2}\int_a^b ds\, h_{\mu\nu}(sp)p_i^\mu p_i^\nu\right], \tag{4.46}$$

where once again s parameterises the contour and $h_{\mu\nu}$ describes the perturbation around a flat background metric as in eq. (4.40). Due to the comparable role of these operators to the Wilson lines in gauge theory, eq. (4.46) has been referred to as a gravitational Wilson line. Further motivation for this identification comes from the double copy.

4.2 A double copy for Wilson lines

4.2.1 Double copy replacements

The similarities between the gauge theory and gravitational Wilson lines in the description of certain physical phenomena suggests that there may be some deeper underlying relation between them. In particular, one might hope that a double copy relation could be set up to provide a more explicit map between the two operators. In ref. [1], it was argued that this is indeed the case.

Let us start with the gauge theory Wilson line of eq. (4.33), written in the form

$$\Phi_A(\gamma) = \mathcal{P}\exp\left[ig\int_\gamma ds A_\mu^a \tilde{T}^a \dot{x}^\mu\right]. \tag{4.47}$$

Here we have temporarily adopted Hermitian generators for the gauge group, such that $\tilde{T}^a = iT^a$. Consider now making the replacements

$$g \to \frac{\kappa}{2}, \quad \tilde{T}^a \to \dot{x}^\mu, \quad A_\mu^a \to h_{\mu\nu}. \tag{4.48}$$

The first is the standard replacement of couplings that occurs in all examples of the double copy. The second is a replacement of colour information with kinematic information, where the latter corresponds to the tangent vector to the Wilson line contour. The final replacement is simply that of the fields themselves. The result of these replacements is the operator

$$\Phi_g(\gamma) = \exp\left[\frac{i\kappa}{2}\int_\gamma ds\, h_{\mu\nu}\dot{x}^\mu \dot{x}^\nu\right], \tag{4.49}$$

which is precisely the gravitational Wilson line, identified in eq. (4.41). One might worry given that this procedure is only valid for a particular choice of einbein in eq. (4.37). However, it is known that in both the BCJ double copy for amplitudes and the classical double copy for exact solutions, the duality is only manifest in certain gauge choices or coordinate systems. As previously mentioned, a choice of einbein corresponds to a gauge choice, and thus the fact that the double copy is manifest for a particular choice is consistent with previous examples of the correspondence.

4.2.2 Relation to scattering amplitudes

Further evidence of this relationship comes from an analysis of the soft factors produced by the two Wilson line operators. As previously mentioned, extracting the soft factor involves taking a VEV of Wilson line operators, with one assigned to each external leg. In practice, this involves including the Wilson line operator in the path integral, such that it appears as an additional contribution to the action. For example, in gauge theory the generating functional from which the soft contributions are generated is [132]

$$Z_s = \int \mathcal{D}A_\mu e^{iS[A_\mu]} \prod_{i=1}^{n} \mathcal{P} e^{-ig\int dx_i^\mu A_\mu(x_i)}, \tag{4.50}$$

Here A_μ is the soft gauge field, $S[A_\mu]$ is the corresponding action, and the product is over the n external legs. Thus, the Wilson line insertions appear in the generating functional as source terms, from which vertices will be generated. To obtain the vertex Feynman rules, we write the Wilson line as in eq. (4.44), and Fourier transform the gauge field,

$$A_\mu(x) = \int \frac{d^dk}{(2\pi)^d} A_\mu(k) e^{ik\cdot x} = \int \frac{d^dk}{(2\pi)^d} A_\mu(k) e^{isk\cdot p}, \tag{4.51}$$

where in the second equality we have inserted the parameterisation of eq. (4.43) for x^μ. The exponent in a given Wilson line is then

$$\begin{aligned} -g\int dx^\mu A_\mu(x) &= -gp^\mu \int_0^\infty ds A_\mu(x) \\ &= -gp^\mu \int_0^\infty ds \int \frac{d^dk}{(2\pi)^d} A_\mu(k) e^{isk\cdot p} \\ &= \int \frac{d^dk}{(2\pi)^d} A_\mu^a(k) \left[g\tilde{T}^a \frac{p^\mu}{k\cdot p}\right], \end{aligned} \tag{4.52}$$

where the integral over s can be formally computed through implementation of the Feynman $i\epsilon$ prescription [130]. The bracketed factor in the final line is the eikonal Feynman rule, which contributes to the appropriate soft factor in gauge theory [132]. It describes the coupling of the soft gauge field to the worldline of the hard particle with momentum p^μ. This analysis can also be performed for the gravitational Wilson line in eq. (4.46), whose exponent yields

$$\frac{i\kappa}{2}\int ds h_{\mu\nu}(x)p^\mu p^\nu = \int \frac{d^dk}{(2\pi)^d}h_{\mu\nu}(k)\left[\frac{\kappa}{2}\frac{p^\mu p^\nu}{k\cdot p}\right] \tag{4.53}$$

The bracketed factor is the eikonal Feynman rule for soft graviton emission [120]. A detailed analysis of the IR-divergent structure of gauge theory and perturbative gravity from the perspective of the double copy was performed in ref. [120]. Here it was concluded that soft factors in each theory are related by the BCJ double copy for amplitudes, thereby providing evidence for its validity at all loop-orders. However, to reach this result ref. [120] carried out a detailed Feynman-diagrammatic analysis. Here we see that it follows straightforwardly from the fact that gauge and gravitational Wilson lines *themselves* exhibit a double copy structure.

4.2.3 Relation to topological patching conditions

The double copy for Wilson lines described here also makes contact with the topological discussion of the previous chapter. Consider the gravitational Wilson line of eq. (4.41), with an integration curve C corresponding to the equator of the two-sphere at infinity. Furthermore, we take $s = mt$ with t the conventional time coordinate. Expanding out the components of the graviton, we then have

$$\Phi_g(C) = \exp\left[\frac{i\kappa m}{2}\oint_C dt\left(h_{00} + 2h_{0i}\dot{x}^i + h_{ij}\dot{x}^i\dot{x}^j\right)\right], \tag{4.54}$$

where $\dot{x}^\mu$ now denotes differentiation with respect to t, so that $\dot{x}^0 = 1$. For asymptotic form of the pure NUT solution of eq. (3.87), we have $h_{00} = 0$. Thus, by considering this solution in the non-relativistic limit, eq. (4.54) reduces to

$$\Phi_g(C) = \exp\left[i\kappa m\oint_C dx^i h_{0i}\right], \tag{4.55}$$

which coincides precisely with eq. (3.91). In the previous chapter, this quantity was used in the construction of a patching condition which classified the non-trivial topology of the pure NUT solution. An analogous patching condition could be set up in gauge theory, also in terms of Wilson lines, to classify the topology of magnetic monopoles, the known single copy of Taub-NUT. Thus, it is not just the case that a topological condition can be given on each side of the double copy correspondence such that it matches the double copy structure of the solutions it classifies, but that this condition is stated in terms of Wilson lines, which themselves exhibit a well-defined double copy.

4.2.4 What is the single copy of the gravitational holonomy?

The discussions of this section and the last raise an interesting question. In section 4.1 we introduced the concept of holonomy, which quantifies the transformation of fields after parallel transport around a closed loop. In both gauge theory and gravity, the holonomy operator is given by an exponentiated integral of a connection. In gauge theory the connection is the gauge field, while in gravity it is either the Christoffel or spin connection. The gauge theory holonomy is therefore a special case of a Wilson line, in which the end points of the curve are taken to be the same point. However, in this section we have argued that the gauge theory Wilson line is the single copy of a quantity, the gravitational Wilson line, which *does not* relate to the gravitational holonomy. Thus, we can conclude that while the gauge and gravitational holonomies are clearly geometric analogues of one another, they cannot be related by the double copy, as the gauge holonomy is already the single copy of a different object. Thus we are left with a question: what is the single copy of the gravitational holonomy?

4.3 The single copy of the gravitational holonomy

The gravitational Wilson line discussed in the previous section contains the action for a spinless point-particle. In order to identify the single copy of the holonomy in gravity, we will need to map this operator to a physical situation whose single copy is already known. As we will see, this turns out to be the dynamics of spinning particles, and so it will be useful to now review the generalisation of point particle actions to include spin.

4.3.1 Relativistic spinning particles

In eq. (4.37) we saw the action for a spinless point particle coupled to gravity. We now want to generalise this to an object with intrinsic angular momentum (for a review see ref. [133]). This could be an extended object such as a black hole, or a point like particle with spin. To begin, we define a vielbein on the worldline, ${e^A}_\mu(\tau)$. The capital latin indices, $\{A, B, C, ...\}$, are those of a body-fixed frame; a frame fixed to the particle as it moves along the worldline such that it spins in tandem with the particle. This vielbein therefore relates the body-fixed frame to a general coordinate frame. It is furthermore related to a general orthonormal frame via a Lorentz transformation:

$$e^A{}_\mu = \Lambda^A{}_a e^a{}_\mu. \tag{4.56}$$

The body-fixed vielbein can be used to define an angular velocity tensor

$$\Omega_{\mu\nu} = e_{A\mu} \frac{De^A{}_\nu}{D\tau}, \tag{4.57}$$

where $D/D\tau$ is the covariant derivative with respect to the worldline parameter, defined as

$$\frac{De^A{}_\nu}{D\tau} \equiv \dot{x}^\alpha D_\alpha e^A{}_\nu = \dot{x}^\alpha \left(\partial_\alpha e^A{}_\nu - \Gamma^\lambda_{\alpha\nu} e^A{}_\lambda \right). \tag{4.58}$$

The angular velocity tensor is antisymmetric. To see this, note that the product rule can be used to rewrite eq. (4.57) as

$$\begin{aligned} \Omega_{\mu\nu} &= \frac{D}{D\tau}(e_{A\mu} e^A{}_\nu) - e^A{}_\nu \frac{De_{A\mu}}{Ds} \\ &= \frac{D}{D\tau}(e_{A\mu} e^A{}_\nu) - \Omega_{\nu\mu}. \end{aligned} \tag{4.59}$$

However, by definition $e_A{}^\mu e^{A\nu} = g^{\mu\nu}$ and thus for a metric connection the first term vanishes to give

$$\Omega_{\mu\nu} = -\Omega_{\nu\mu}. \tag{4.60}$$

The full action for the spinning object is now given by

$$S_{pp} = S^{(0)}_{pp} + S^{(1)}_{pp}, \tag{4.61}$$

where $S^{(0)}_{pp}$ is the spinless point particle action of eq. (4.37), and the correction due to the spin is given by

$$S^{(1)}_{pp} = -\frac{1}{2} \int d\tau \, \Omega_{\mu\nu} S^{\mu\nu}. \tag{4.62}$$

Here $S_{\mu\nu}$ is the spin tensor, defined as the dynamical variable conjugate to the angular momentum. In writing eq. (4.62) we have not included an additional gauge fixing term that is usually present to eliminate residual degrees of freedom in the spin tensor. For our arguments this term will not play a role.

4.3.2 Holonomy from a spinning particle

We will now argue that the dynamics of spinning particle acts as a physical manifestation of the gravitational holonomy operator of eq. (4.31). To this end, we note that in general a given vielbein ${e^a}_\mu$ will not align with the body-fixed vielbein ${e^A}_\mu$. Recall, however, that there will be a Lorentz transformation that will relate the two, as in eq. (4.56). Let us then rewrite the angular velocity tensor in eq. (4.57) as

$$\Omega_{\mu\nu} = S_{\mu\nu} {\Lambda_A}^a {e_a}^\mu \frac{D\Lambda^{Ab} {e_b}^\nu}{D\tau}. \tag{4.63}$$

Expanding out the covariant derivative and using the known relation between the Christoffel and spin connections,

$$\Gamma^\sigma_{\mu\nu} = {e_a}^\sigma (\omega_\mu)^a{}_b {e^b}_\nu + {e_a}^\sigma \partial_\mu {e^a}_\nu, \tag{4.64}$$

the angular velocity tensor reduces to

$$\Omega_{\mu\nu} = {e^a}_\mu {e^b}_\nu \left(\Lambda_{Aa} {\dot{\Lambda}^A}{}_b - (\omega_\rho)_{ab} \dot{x}^\rho \right). \tag{4.65}$$

The contraction of the angular velocity and spin tensors that appears in the spin correction to the point particle action in eq. (4.62) can then be written as

$$\Omega_{\mu\nu} S^{\mu\nu} = S^{ab} \left(\Lambda_{Aa} {\dot{\Lambda}^A}{}_b - (\omega_\rho)_{ab} \dot{x}^\rho \right). \tag{4.66}$$

The two terms in this expression have simple physical interpretations. The action of eq. (4.62) describes the dynamics of the spin of the object, or in other words how the body-fixed vielbein changes as we proceed along the worldline. The action therefore governs how a vector fixed to the moving object transforms with the motion of the object. There are two contributions to this change: the rotation of the object and the fact that the body fixed vielbein is changing due to the underlying spacetime. These two contributions correspond precisely to the first and second terms of eq. (4.66) respectively. This can be seen by noticing that the first term contains the flat-space definition of the angular velocity tensor

$$\Omega^{\rm flat}_{ab} = \Lambda_{Aa} \frac{d{\Lambda^A}_b}{d\tau}, \tag{4.67}$$

and thus would be present even in a trivial background spacetime. The second term in eq. (4.66) then governs the transformation of a vector fixed to the object due to motion in a non-trivial background.

Recall that the spinless point-particle action can be exponentiated to give a gravitational Wilson line operator, representing the phase picked up by a particle as it traverses a contour. That is, the gravitational Wilson line is given by

$$\Phi_g = e^{iS^{(0)}_{pp}}, \tag{4.68}$$

where $S^{(0)}_{pp}$ is given in eq. (4.37), and one discards terms associated with flat space, which amount to multiplicative factors that vanish for appropriately normalised VEVs of Wilson line operators. Following the discussion of the previous section, we can now form an object corresponding to the phase experienced by a spinning particle by considering

$$\Phi^{\text{spin}}_g = e^{iS_{pp}}, \tag{4.69}$$

where S_{pp} is now the spinning point-particle action of eq. (4.61). The spin tensor S^{ab} is valued in the algebra of the Lorentz group, such that we may write

$$S^{ab}(\tau) = Q^{ab}_{cd}(\tau)M^{cd} \tag{4.70}$$

where the $\{M^{ab}\}$ are the Lorentz generators. The quantity $Q^{ab}_{cd}(\tau)$ dictates the relative "strength" of each generator as the object traverses a given contour. We are interested in the most general possible case, as we wish to examine how a general vector is modified after transport around a general loop. Thus, we choose a spin tensor such that

$$Q^{ab}_{cd} = \frac{1}{2}\left(\delta^a_c\delta^b_d - \delta^b_c\delta^a_d\right), \tag{4.71}$$

which physically amounts to a democratic assignment of unit spin along all axes. In the case of a quantum particle in state $|\psi\rangle$, the spin tensor is given by a normalised expectation value

$$S^{ab} = \frac{\langle\psi|Q^{ab}_{cd}M^{cd}|\psi\rangle}{\langle\psi|\psi\rangle}, \tag{4.72}$$

however one may still make the choice of eq. (4.71). In practical terms, this amounts the replacement,

$$S^{ab} \to M^{ab}, \tag{4.73}$$

in the definition of our spin-corrected Wilson line in eq. (4.69). The result is then

$$\Phi^{\text{spin}}_g(\gamma) = \mathcal{P}\exp\left[\frac{i\kappa}{2}\int_\gamma ds\left(h_{\mu\nu}\dot{x}^\mu\dot{x}^\nu - \dot{x}^\mu(\omega_\mu)_{ab}M^{ab}\right)\right], \tag{4.74}$$

where the path ordering is necessary due to the matrix-valued Lorentz generators in the second term. Note also that there is an implicit identity matrix in the first term. This object describes the phase experienced by a particle if it has spin, with the second term in the exponential dictating how the spin-dependent degrees-of-freedom couple to a non-trivial gravitational field.

The second term in the exponential of eq. (4.74) is precisely the gravitational holonomy operator of eq. (4.31), prior to taking the integration curve to be a closed loop. Here we see that the natural geometrical counterpart to the gauge theory Wilson line, the exponentiated integral of the spin connection, appears as a spin-dependent correction to the gravitational Wilson line that is identified via the double copy. It is interesting to note that a similar observation was made as early as the 1960s [134], prior to the introduction of Wilson lines.

4.3.3 The single copy of the holonomy

We have now seen that the Wilson line constructed from the spinning point-particle action contains the gravitational holonomy operator. From this observation we can already see how to identify a single copy of this operator. We simply consider the known action of a spinning particle coupled to a gauge field, and from this construct a generalised Wilson line containing a spin-correction to the phase. The action for a spinning point-particle coupled to a gauge field is (see e.g. ref. [135])

$$S_{\text{gauge}} = \int d\tau \left[\frac{1}{2e(\tau)} \eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu - \frac{e(\tau)m^2}{2} + \frac{1}{2} \Omega^{\text{flat}}_{\mu\nu} S^{\mu\nu} \right.$$
$$\left. + \, g c^a(\tau) \left(\dot{x}^\mu A^a_\mu - \frac{e(\tau)}{2} F^a_{\mu\nu} S^{\mu\nu} \right) \right] . \tag{4.75}$$

Here the first two terms make up the flat space spinless point-particle action. The third term acts as the spin correction to this, and thus corresponds to the first term in eq. (4.66). The remaining terms represent the coupling of the particle to a gauge field, where $c^a(\tau)$ is a colour vector obtained at a given point along the worldline by the expectation value of the colour generator $\tilde{T}^a$. The first of these terms corresponds to the standard gauge theory Wilson line, once the colour vector is replaced with the generator itself. The final term contains a contraction of the field strength $F^a_{\mu\nu}$ with the spin tensor. It therefore represents a spin-dependent correction to the vacuum dynamics of a spinning particle, that arises due to the presence of a non-trivial gauge field background. This term is thus the gauge theory analogue of the spin connection term in eq. (4.66).

To clarify these statements, we can furnish the spin-dependent correction with a nice physical interpretation. We focus here on the case of abelian gauge theory for simplicity.

We can identify the electric and magnetic fields respectively as

$$F_{0i} = E_i, \qquad F_{ij} = \epsilon_{ijk} B_k, \tag{4.76}$$

such that the contraction of the field strength and spin tensor in the final term of eq. (4.75) is

$$\begin{aligned} F_{\mu} S^{\mu\nu} &= 2F_{0i} S^{0i} + F_{ij} S^{ij} \\ &= -2(E_i d_i + B_i \mu_i). \end{aligned} \tag{4.77}$$

Here we have defined the electric dipole moment $\boldsymbol{d}$ and the magnetic dipole moment $\boldsymbol{\mu}$, such that

$$d_i = -S_{0i}, \qquad \mu_i = -\frac{1}{2} \epsilon_{ijk} S^{jk}. \tag{4.78}$$

Equation (4.77) therefore represents the coupling of the test particle to the gauge field when both the electric and magnetic dipole moments are turned on in general.

The fact that the spinning point-particle actions in gauge theory and gravity can be related via the double copy has been addressed in great detail in refs. [59, 135–137]. These works carried out a perturbative analysis of the radiation emitted by spinning classical sources described by the action of eq. (4.75), and double copied the results order-by-order in perturbation theory. The double copy was found to correspond to the radiation emitted by a particle interacting with a graviton, dilaton, and axion. This is to be expected, as the double copy of pure Yang-Mills is $\mathcal{N} = 0$ supergravity, rather than pure general relativity. One might then ask about the absence of these fields in our discussion of holonomy. Here we wish to identify the operator that is the single copy of the gravitational holonomy, a quantity defined in pure general relativity. As the single copy of pure Einstein gravity or $\mathcal{N} = 0$ supergravity is always pure Yang-Mills theory, we can unambiguously identify the single copy of the graviton spin coupling as the final term in eq. (4.75).

Following this discussion, we find the generalised Wilson line in gauge theory to be

$$\Phi_{\text{spin}}(\gamma) = \mathcal{P} \exp \left[ig \int_{\gamma} ds \left(A^a_{\mu} \dot{x}^{\mu} - \frac{1}{2} F^a_{\mu\nu} M^{\mu\nu} \right) \tilde{T}^a \right], \tag{4.79}$$

where we have fixed the einbein to be $e = 1$. The first term in the exponential corresponds to the standard gauge theory Wilson line, while the second represents the correction to this due to spin. Evaluating the spin correction over a closed curve C

yields an operator that we identify as the single copy of the gravitational holonomy:

$$\Phi_F(C) = \mathcal{P}\exp\left[-\frac{ig}{2}\oint_C ds\,\tilde{T}^a F^a_{\mu\nu} M^{\mu\nu}\right]. \tag{4.80}$$

Here we have motivated the identification of this operator as the single copy of the gravitational holonomy by noting its relation to spinning point particle actions and their double copy structure. We will now provide further evidence of this identification by studying the role of the operator in the context of the soft behaviour of amplitudes and exact Kerr-Schild solutions.

4.3.4 Relation to amplitudes

The identification of eq (4.80) as the single copy of the gravitational holonomy operator provides a nice link to the well known next-to-soft theorems for the emission of low energy radiation [138, 139]. Let us first consider the gravity case, where the relevant operator is the Riemannian parallel propagator of eq. (4.3) (recall that this becomes the holonomy when the integration contour is a closed loop). We now write this in terms of the metric by using the explicit form of the Christoffel symbols,

$$\Gamma^\mu_{\rho\sigma} = \frac{1}{2}g^{\mu\alpha}\left(\partial_\rho g_{\alpha\sigma} + \partial_\sigma g_{\alpha\rho} - \partial_\alpha g_{\rho\sigma}\right), \tag{4.81}$$

and consider perturbations around the flat metric as in eq. (4.40). The result is the operator

$$[\Phi_\Gamma(\gamma)]^\mu{}_\sigma = \mathcal{P}\exp\left[-\frac{\kappa}{2}\int_\gamma dx^\rho\left(\partial_\rho h^\mu_\sigma + \partial_\sigma h^\mu_\rho - \partial^\mu h_{\rho\sigma} + \ldots\right)\right], \tag{4.82}$$

where the ellipsis denotes higher orders in κ. For the integration contour we choose a straight line out to infinity, originating at the origin. This is appropriate for a fast-moving particle leaving a given scattering process, and allows us to parameterise x via

$$x^\mu = sp^\mu, \quad 0 \leq s \leq \infty. \tag{4.83}$$

The first term in eq. (4.82) is a total derivative. It will integrate to give a gauge-dependent artifact associated with the endpoints of the integration contour, and will vanish when computing gauge invariant amplitudes.[1] By ignoring this term and intro-

[1]This assumption may not hold in the computation of certain observables, in which momentum derivatives act on the parallel propagator.

ducing Fourier components for the graviton field via

$$h_{\mu\nu}(x) = \int \frac{d^d k}{(2\pi)^d} \tilde{h}_{\mu\nu}(k) e^{-ik\cdot x}, \tag{4.84}$$

the remaining terms take the form

$$-\frac{\kappa}{2}\int_0^\infty ds p^\rho \left(\partial_\sigma h^\mu_\rho - \partial^\mu h_{\rho\sigma}\right) = \frac{i\kappa}{2} p^\rho \int \frac{d^d k}{(2\pi)^d} \int_0^\infty ds \left(k_\sigma \tilde{h}^\mu_\rho - k^\mu \tilde{h}_{\rho\sigma}\right) e^{-isk\cdot p}. \tag{4.85}$$

Performing the s integral then yields

$$\ln(\Phi_g) \sim \int \frac{d^d k}{(2\pi)^d} \tilde{h}_{\beta\rho}(k) \left[\frac{\kappa}{2} \frac{p^\rho k_\alpha (M^{\alpha\beta})^\mu{}_\sigma}{p\cdot k}\right], \tag{4.86}$$

where we have introduced the spin-1 Lorentz generators of eq. (4.30). The factor in the square brackets can be recognised as the appropriate contribution to the next-to-soft theorem for graviton emission [138]. Here it arises as a correction to the soft factor generated by the gravitational Wilson line of eq. (4.41), representing the spin-dependent coupling to the hard particle worldline. This can be seen as it is suppressed by a single power of the momentum of the emitted radiation. If we perform a similar analysis in gauge theory, using the single copy operator of eq. (4.80), we find

$$\ln(\Phi_F) \sim \int \frac{d^d k}{(2\pi)^d} \tilde{A}^a_\mu(k) \left[g \tilde{T}^a \frac{k_\nu M^{\mu\nu}}{p\cdot k}\right]. \tag{4.87}$$

The bracketed factor is now in agreement with the appropriate next-to-soft theorem in gauge theory [139]. The known double copy properties of the next-to-soft factors therefore provides further evidence for the identification of eq. (4.80) as the single copy of the gravitational holonomy operator [119, 140, 141].

4.3.5 Insights from Kerr-Schild solutions

Let us now specialise to the case of Kerr-Schild solutions, for which the exact double copy properties are well known [21]. Recall that Kerr-Schild solutions in general relativity take the exact form

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}, \tag{4.88}$$

where the graviton field $h_{\mu\nu}$ decomposes such that

$$h_{\mu\nu} = \phi k_\mu k_\nu. \tag{4.89}$$

The Kerr-Schild vectors k_μ satisfy the null and geodesic conditions given in eq. (2.62), and we have taken the background metric in eq. (4.88) to be Minkowski spacetime. The single copy of the graviton field is then a gauge field which takes the form

$$A^a_\mu = \phi c^a k_\mu, \tag{4.90}$$

and satisfies the Yang-Mills equations, which linearise on this solution.

We will now study the form of the gravitational holonomy for Kerr-Schild metrics. In an orthonormal basis, one has

$$g_{\mu\nu} = \eta_{ab} e^a{}_\mu e^b{}_\nu, \tag{4.91}$$

which in turn implies the following form for the Kerr-Schild vielbein:

$$e^a{}_\mu = \bar{e}^a{}_\mu + \frac{1}{2}\phi k^a k_\mu, \quad e_a{}^\mu = \bar{e}_a{}^\mu - \frac{1}{2}\phi k_a k^\mu. \tag{4.92}$$

Here $\bar{e}^a{}_\mu$ is the vielbein associated with the background metric in eq. (4.88), which for our purposes is the Minkowski metric $\eta_{\mu\nu}$. We now want to derive the form of the spin connection in Kerr-Schild coordinates. This is done explicitly in Appendix A, where the result is found to be

$$(\omega_\mu)_{ab} = \partial_b e_{a\mu} - \partial_a e_{b\mu}. \tag{4.93}$$

Unlike the general expression of eq. (4.22), this has the pleasing property of being linear in the vielbein. Substituting the results of eq. (4.92), with indices lowered appropriately, into eq. (4.93) yields

$$(\omega_\mu)_{ab} = \frac{1}{2}\left[\partial_b(\phi k_\mu k_a) - \partial_a(\phi k_\mu k_b)\right] \tag{4.94}$$

$$= \frac{1}{2}\left[e_b{}^\sigma \partial_\sigma(\phi k_\mu k_a) - e_a{}^\sigma \partial_\sigma(\phi k_\mu k_b)\right]. \tag{4.95}$$

Note that due to the null property of the Kerr-Schild vectors, $k^\mu k_\mu = 0$, conversion between coordinate and orthonormal bases is done simply with the background vielbein:

$$k_a = e_a{}^\mu k_\mu = \bar{e}_a{}^\mu k_\mu - \frac{1}{2}\phi k_a k^\mu k_\mu = \bar{e}_a{}^\mu k_\mu. \tag{4.96}$$

Thus, the spin connection in eq. (4.95) can be written as

$$(\omega_\mu)_{ab} = \frac{1}{2}\left[\bar{e}_a{}^\nu e_b{}^\sigma - \bar{e}_b{}^\nu e_a{}^\sigma\right] \partial_\sigma(\phi k_\mu k_\nu). \tag{4.97}$$

If we now write the remaining vielbeins explicitly, a great deal of simplification occurs.

To see this, consider the first term in the above expression:

$$\bar{e}_a{}^\nu e_b{}^\sigma \partial_\sigma(\phi k_\mu k_\nu) = \bar{e}_a{}^\nu \left[\bar{e}_b{}^\sigma - \frac{1}{2}\phi k_b k^\sigma\right] \partial_\sigma(\phi k_\mu k_\nu) \tag{4.98}$$

$$= \bar{e}_a{}^\nu \bar{e}_b{}^\sigma \partial_\sigma(\phi k_\mu k_\nu) - \frac{1}{2}\phi k_a k_b k_\mu k^\sigma \partial_\sigma \phi. \tag{4.99}$$

In the second equality, we have expanded $\partial_\sigma(\phi k_\mu k_\nu)$ using the product rule, from which two of the three resulting terms vanish due to the geodesic condition $k^\sigma \partial_\sigma k_\mu = 0$. Performing the same procedure for the second term in eq. (4.97), we find that the terms which contain only a derivative of the scalar field ϕ cancel, such that eq. (4.97) is simply

$$(\omega_\mu)_{ab} = \frac{1}{2}\left[\bar{e}_a{}^\nu \bar{e}_b{}^\sigma - \bar{e}_b{}^\nu \bar{e}_a{}^\sigma\right] \partial_\sigma(\phi k_\mu k_\nu). \tag{4.100}$$

Finally, if we expand the spin connection in terms of Lorentz generators, we obtain

$$\frac{i}{2}(\omega_\mu)_{cd} M^{cd} = -\frac{i}{2}\partial_\sigma(\phi k_\mu k_\nu) M^{\nu\sigma}, \tag{4.101}$$

where we have identified the spin-1 Lorentz generators as

$$(M^{\nu\sigma})_{ab} = i\left[\bar{e}_a{}^\nu \bar{e}_b{}^\sigma - \bar{e}_b{}^\nu \bar{e}_a{}^\sigma\right]. \tag{4.102}$$

Thus, we are left with a simple expression in which the exponent appearing in the Kerr-Schild gravitational holonomy operator is written directly in terms of the graviton:

$$\oint dx^\mu (\omega_\mu)_{ab} M^{ab} = -\oint dx^\mu \partial_\sigma(h_{\mu\nu}) M^{\nu\sigma}. \tag{4.103}$$

We now single copy this expression by making the replacements

$$\dot{x}^\mu \to \tilde{T}^a, \quad k_\mu \to c^a, \tag{4.104}$$

in line with the Kerr-Schild single copy. The result is then

$$\oint dx^\mu (\omega_\mu)_{ab} M^{ab} \to -\tilde{T}^a \oint ds\, \partial_\sigma(\phi k_\nu c^a) M^{\nu\sigma} = -\tilde{T}^a \oint ds\, \partial_\sigma(A^a_\nu) M^{\nu\sigma} \tag{4.105}$$

$$= \frac{1}{2}\tilde{T}^a \oint ds\, F^a_{\nu\sigma} M^{\nu\sigma}. \tag{4.106}$$

This agrees with the conclusion reached above, that the single copy of the gravitational holonomy is the operator of eq. (4.80).

Gauge Theory	Gravity
$\mathcal{P}\exp\left[-g\oint_C ds\,\dot{x}^\mu A_\mu\right]$	$\exp\left[\frac{i\kappa}{2}\oint_C ds\,\dot{x}^\mu\dot{x}^\nu h_{\mu\nu}\right]$
$\mathcal{P}\exp\left[\frac{g}{2}\oint_C ds\,F_{\mu\nu}M^{\mu\nu}\right]$	$\mathcal{P}\exp\left[-\frac{i\kappa}{2}\oint_C ds\,\dot{x}^\mu(\omega_\mu)_{ab}M^{ab}\right]$

Table 4.1: Wilson line and holonomy operators in gauge and gravity theories, and their single / double copies.

4.3.6 A square of four operators

To summarise the results of this section and the last, we can construct a square of four operators that are related via the double copy, as shown in table 4.1. The holonomy in gauge theory and gravity appears in the top-left and bottom-right respectively, while the gravitational Wilson line is in the top-right. In section 4.2, we introduced a double copy relating the operators in the top row of the table, namely Wilson lines in gauge theory and gravity. Previous works have noted the analogous physical roles played by these operators, and here we have shown that this can be interpreted as an instance of the double copy. Furthermore, in this section we have introduced a new operator, that in the bottom-left of the table, and interpreted it as the single copy of the gravitational holonomy. This operator has not previously been related to the gravitational holonomy, due to the fact that it does not arise from an analogous geometric definition. In fact, its role is entirely different to the holonomy in gauge theory. The standard holonomy describes how vectors in the internal colour space are transformed as a result of parallel transport. In contrast, the single copy of the gravitational holonomy acts as a transformation on both colour and spacetime vectors, thereby linking the gauge field with vectors in the tangent space of the base manifold.

Now that we have identified eq. (4.80) as the single copy of the gravitational holonomy, it is natural to ask if this operator has any use. We explore this question in the next section.

4.4 The single copy holonomy operator

Beyond its physical role in the description of how vectors are transformed after parallel transport, the gravitational holonomy is also used in the classification of solutions in general relativity and arbitrary manifolds more generally. The maximal holonomy group is $\mathrm{GL}(m)$, where m is the dimension of the manifold. For a metric connection on an orientable manifold, this group reduces to $\mathrm{SO}(m)$ for a Riemannian manifold or $\mathrm{SO}(m-1,1)$ for a Lorentzian manifold. Further reductions occur as we consider manifolds with other special properties. The classification of manifolds based on how the holonomy reduces has been widely studied, leading, in the most well-known case, to Berger's classification [123]. It is thus natural to ask whether the operator we have identified as the single copy of the gravitational holonomy has a similar purpose. In the following, for ease of reference, we will refer to eq. (4.80) as the *SCH operator*, short for "single copy holonomy operator". Furthermore, we will refer to the group of transformations generated by this operator as the *SCH group*. We now want to ask whether the SCH group reduces in certain cases, or more interestingly, whether the same reduction occurs for the SCH and holonomy groups when they correspond to solutions related by the double copy. We will investigate this by considering three well-known instances of exact double copies.

4.4.1 The Schwarzschild solution

We begin by considering the Schwarzschild solution in general relativity. This is sourced by a point mass M at the origin, and has a known Kerr-Schild form. Working in spherical coordinates (t,r,θ,φ), the scalar field and Kerr-Schild vector take the form[2]

$$\phi(r) = \frac{M}{4\pi r}, \quad k_\mu = (1,1,0,0). \tag{4.107}$$

These enter the definition of the full metric as in eq. (4.88). In section 4.3.5 we derived a form for the spin connection in Kerr-Schild coordinates, which we once again make use of here. In particular, eq. (4.103) gives the Kerr-Schild form for the factor appearing in the exponential of the gravitational holonomy. We now choose a selection of integration contours C and see what elements of the holonomy group we find. Let us first choose a circular orbit in the equatorial plane, with t fixed. We can parameterise this with

$$C: \quad x^\mu = (0,0,0,\varphi), \quad \varphi \in [0,2\pi). \tag{4.108}$$

[2]Beware the use of both φ and ϕ in this section. ϕ is the scalar field appearing in the Kerr-Schild form of the metric, while φ is the azimuthal angle in spherical coordinates.

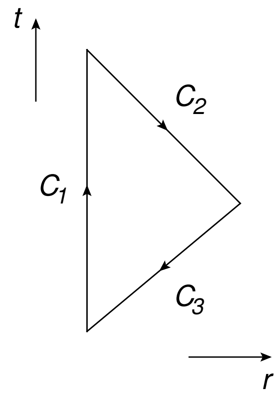


Figure 4.1: A loop formed from three segments in the (r, t) plane. The segments C_2 and C_3 are null lines.

Using the relation in eq. (4.103), we compute the integral appearing in the holonomy to be

$$\oint_C d\varphi\, \partial_\sigma(h_{\varphi\nu}) M^{\nu\sigma} = 0, \tag{4.109}$$

where we have used the fact that, according to eq. (4.107), the Kerr-Schild graviton has no non-zero $h_{\varphi\nu}$ components. The element of the holonomy group associated with this curve is therefore the identity element. We can generate new constant time curves by tilting C with respect to the equatorial plane. However, spherical symmetry implies that these loops will also correspond to a trivial holonomy.

We can, however, generate a non-trivial holonomy for the Schwarzschild solution, via a careful choice of curve. Consider the loop depicted in figure 4.1. This consists of three segments in the (r, t) plane. The first segment C_1 is parallel to the time axis, and can be parameterised via

$$C_1 : \quad x^\mu = (t, r_0, 0, 0), \quad 0 \leq t \leq T. \tag{4.110}$$

This segment is at a fixed radius $r = r_0$ and has total length T. We have also fixed the remaining coordinates such that $\theta = \varphi = 0$. Let us now again make use of eq. (4.103), with which we find

$$\begin{aligned} \int_{C_1} dx^\mu \partial_\sigma(h_{\mu\nu}) M^{\nu\sigma} &= \int_0^T dt \partial_r(h_{00}) M^{0r} \\ &= -\frac{MT}{4\pi r_0^2} M^{0r}. \end{aligned} \tag{4.111}$$

Turning to the other line segments in figure 4.1, we have the following parameterisations

$$C_2: \quad x^\mu = (t, r_0 + T - t, 0, 0), \quad T \geq t \geq T/2, \tag{4.112}$$

$$C_3: \quad x^\mu = (t, r_0 + t, 0, 0)\,, \quad \frac{T}{2} \geq t \geq 0. \tag{4.113}$$

Performing the integrals, one finds

$$\int_{C_2} dx^\mu \partial_\sigma (h_{\mu\nu}) M^{\nu\sigma} = \int_T^{T/2} dt \partial_r (h_{00}) M^{0r}\Big|_{r=r_0+T-t} + \int_{r_0}^{r_0+T/2} dr \partial_r (h_{r0}) M^{0r}; \tag{4.114}$$

$$\int_{C_3} dx^\mu \partial_\sigma (h_{\mu\nu}) M^{\nu\sigma} = \int_{T/2}^{0} dt \partial_r (h_{00}) M^{0r}\Big|_{r=r_0+t} + \int_{r_0+T/2}^{r_0} dr \partial_r (h_{r0}) M^{0r}. \tag{4.115}$$

In summing these terms, the radial integrals cancel, such that after using eq. (4.107) we have

$$\int_{C_2 \cup C_3} dx^\mu \partial_\sigma (h_{\mu\nu}) M^{\nu\sigma} = \frac{M}{2\pi} \frac{T}{r_0 (T + 2r_0)} M^{0r}. \tag{4.116}$$

Finally, we can combine this with the C_1 contribution from eq. (4.111) to give the full contribution from the entire loop:

$$\oint dx^\mu \partial_\sigma (h_{\mu\nu}) M^{\nu\sigma} = \alpha M^{0r}, \qquad \alpha = -\frac{M}{4\pi} \frac{T^2}{r_0^2 (T + 2r_0)}. \tag{4.117}$$

This corresponds to an infinitesimal boost in the (r, t) plane with hyperbolic angle α. To see this, recall that we can define the infinitesimal generators for boots K_i and rotations J_i in terms of the $\{M_{ij}\}$ via

$$K_i = M^{0i}, \quad J_i = \frac{1}{2} \epsilon_{ijk} M^{jk}, \tag{4.118}$$

In terms of K_i and J_i, the Lorentz algebra

$$[M^{\mu\nu}, M^{\rho\sigma}] = i \left(\eta^{\sigma\mu} M^{\rho\nu} + \eta^{\nu\sigma} M^{\mu\rho} - \eta^{\rho\mu} M^{\sigma\nu} - \eta^{\nu\rho} M^{\mu\sigma} \right) \tag{4.119}$$

can be written as

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \quad [J_i, K_j] = i\epsilon_{ijk} K_k, \quad [K_i, K_j] = -i\epsilon_{ijk} J_k. \tag{4.120}$$

Clearly, eq. (4.117) contains the boost generator K_r. By considering loops with fixed r_0 but varying ϕ and φ, we can generate the full set of boosts associated with arbitrary directions. Furthermore, note that boosts do not form a closed algebra, and thus by exponentiating the boost generators we will produce generic transformations corresponding to arbitrary combinations of boosts and rotations. Thus, we conclude that the holonomy group of Schwarzschild spacetime is SO$(d-1,1)$. Note that this is in agreement with the conclusion reached by ref. [142]. Our explicit results differ as we have considered a different orientation for the loop, and made use of Kerr-Schild coordinates. However, our conclusions are qualitatively equivalent.

Let us now turn to gauge theory. The single copy of the Schwarzschild solution is well known to be an abelian-like point charge in Yang-Mills theory [21]. For this solution, the only non-zero component of the field strength is

$$F_{0r} = \frac{Q_E}{4\pi r^2}, \tag{4.121}$$

where Q_E is the electric charge. Consider now the exponent of the SCH operator, which we can take to be abelian due to the nature of the gauge theory solution

$$\ln(\Phi_F) = -\frac{ig}{2}\oint_C ds F_{\mu\nu} M^{\mu\nu}. \tag{4.122}$$

Clearly, in plugging eq. (4.121) into this expression we find an infinitesimal boost in the r, t plane. Thus, by the same arguments given in the gravity case, the SCH group for the point charge is SO$(d-1,1)$. In this case, we have therefore found that two solutions related via the classical double copy are classified by the same holonomy and SCH groups. It is reasonable to wonder whether this should always be the case. By considering a slightly more complex example, we will now see that it is not.

4.4.2 The Taub-NUT solution

The Taub-NUT solution in general relativity is non-asymptotically flat metric characterised by a Schwarzschild like mass M and a NUT charge N. The properties of Taub-NUT spacetime were discussed in the previous chapter, in section 3.1.2. For our purposes here, it will be useful to write the metric in the form

$$ds^2 = -A(r)\left[dt + B(\theta)d\phi\right]^2 + A^{-1}(r)dr^2 + C(r)\left[d\theta^2 + D(\theta)d\phi^2\right], \tag{4.123}$$

where for convenience we have defined

$$A(r) = \frac{(r-r_+)(r-r_-)}{r^2+N^2}, \quad B(\theta) = 2N\cos\theta, \quad C(r) = r^2 + N^2, \quad D(\theta) = \sin^2\theta, \tag{4.124}$$

and

$$r_\pm = M \pm \sqrt{M^2+N^2}. \tag{4.125}$$

As reviewed in the previous chapter, the single copy of the Taub-NUT solution is an electromagnetic dyon, where the mass maps to electric charge and the NUT charge maps to magnetic charge. Here we wish to examine the relationship between the SCH group in gauge theory and the holonomy group in gravity. In the $N \to 0$ limit, the Taub-NUT metric reduces to the Schwarzschild solution. As we have already examined this case in the previous section, we instead consider here the $M \to 0$ limit, corresponding to the pure NUT solution. For simplicity, we will consider the abelian case. However, as detailed in the previous chapter, the generalisation to the non-abelian case is straightforward.

For an abelian magnetic monopole, the non-zero components of the field strength are

$$F_{\theta\phi} = \frac{Q_M}{4\pi r^2}, \tag{4.126}$$

where Q_M is the magnetic charge. Thus, the exponent of the SCH operator is simply

$$\ln(\Phi_F) = -\frac{ig}{2}\oint_C ds \frac{Q_M}{4\pi r^2} M^{\theta\phi}, \tag{4.127}$$

such that only the generator for rotations in the (θ, ϕ) plane are turned on. This integral is non-zero in general. Consider, for example, a constant time curve in the equatorial plane at fixed radius $r = r_0$. All factors in eq. (4.127) can then be taken outside of the integral, such that upon performing the integration we simply obtain the length of the curve. Now varying θ and ϕ we will generate all possible rotations, but not boosts. Returning to eq. (4.120) we see that the rotation algebra closes upon itself, such that exponentiation of eq. (4.127) yields only rotations. Thus, we conclude that the SCH group of the magnetic monopole solution in gauge theory is SO(3), and is therefore reduced compared to the electric case for which we obtained SO(3, 1).

Let us now consider how this compares to the holonomy group of the Taub-NUT solution in gravity. Recall that a vielbein basis is chosen such that

$$g = g_{\mu\nu} dx^\mu \otimes dx^\nu = \eta_{ab} dx^a \otimes dx^b. \tag{4.128}$$

Thus, a natural basis for the Taub-NUT metric in eq. (4.123) is

$$e^0 = A^{\frac{1}{2}}(dt + Bd\phi), \quad e^1 = A^{-\frac{1}{2}} dr, \quad e^2 = C^{\frac{1}{2}} d\theta, \quad e^3 = (CD)^{\frac{1}{2}} d\phi. \tag{4.129}$$

It will also be useful to have the inverse of these expressions:

$$dt = A^{-\frac{1}{2}} e^0 - B(CD)^{-\frac{1}{2}} e^3, \quad dr = A^{\frac{1}{2}} e^1, \quad d\theta = C^{-\frac{1}{2}} e^2, \quad d\phi = (CD)^{-\frac{1}{2}} e^3. \tag{4.130}$$

The spin connection can be obtained from the torsion-free form of Cartan's structure equations,

$$\omega^a{}_b \wedge e^b = -de^a, \tag{4.131}$$

along with the metric compatibility condition

$$\omega_{ab} = -\omega_{ba}. \tag{4.132}$$

Thus, we first calculate the exterior derivatives of the basis in eq. (4.129), and use eq. (4.130) to write the results in terms of the vielbein basis:

$$de^0 = (\partial_r A^{\frac{1}{2}}) e^1 \wedge e^0 + (\partial_\theta B) C^{-1} \left(\frac{A}{D}\right)^{\frac{1}{2}} e^2 \wedge e^3, \tag{4.133}$$

$$de^1 = 0, \tag{4.134}$$

$$de^2 = (\partial_r C^{\frac{1}{2}}) \left(\frac{A}{C}\right)^{\frac{1}{2}} e^1 \wedge e^2, \tag{4.135}$$

$$de^3 = (\partial_r C^{\frac{1}{2}}) \left(\frac{A}{C}\right)^{\frac{1}{2}} e^1 \wedge e^3 + (\partial_\theta D^{\frac{1}{2}}) \left(\frac{1}{CD}\right)^{\frac{1}{2}} e^2 \wedge e^3. \tag{4.136}$$

These expressions can now be used in the Cartan structure equations of eq. (4.131) which, along with the metric compatibility condition in eq. (4.132), yields the following non-zero components of the spin connection:

$$\omega^0{}_1 = \omega^1{}_0 = (\partial_r A^{\frac{1}{2}}) A^{\frac{1}{2}} (dt + B d\phi), \tag{4.137}$$

$$\omega^0{}_2 = \omega^2{}_0 = \frac{1}{2} (\partial_\theta B) \left(\frac{A}{C}\right)^{\frac{1}{2}} d\phi, \tag{4.138}$$

$$\omega^0{}_3 = \omega^3{}_0 = -\frac{1}{2} (\partial_\theta B) \left(\frac{A}{CD}\right)^{\frac{1}{2}} d\theta, \tag{4.139}$$

$$\omega^1{}_2 = -\omega^2{}_1 = -(\partial_r C^{\frac{1}{2}}) A^{\frac{1}{2}} d\theta, \tag{4.140}$$

$$\omega^1{}_3 = -\omega^3{}_1 = -(\partial_r C^{\frac{1}{2}}) (AD)^{\frac{1}{2}} d\phi \tag{4.141}$$

$$\omega^2{}_3 = -\omega^3{}_2 = -\frac{1}{2} (\partial_\theta B) \frac{A}{C D^{\frac{1}{2}}} (dt + B d\phi) - (\partial_\theta D^{\frac{1}{2}}) d\phi. \tag{4.142}$$

Now by substituting eqs. (4.124) and performing the derivatives we find for Taub-NUT:

$$\omega^0{}_1 = \omega^1{}_0 = \frac{M(r^2 - N^2) + 2N^2 r}{(r^2+N^2)^2}\left[dt + 2N\cos\theta d\phi\right], \tag{4.143}$$

$$\omega^0{}_2 = \omega^2{}_0 = -\frac{N\sin\theta}{r^2+N^2}\sqrt{(r-r_+)(r-r_-)}d\phi, \tag{4.144}$$

$$\omega^0{}_3 = \omega^3{}_0 = \frac{N}{r^2+N^2}\sqrt{(r-r_+)(r-r_-)}d\theta, \tag{4.145}$$

$$\omega^1{}_2 = -\omega^2{}_1 = -\frac{r}{r^2+N^2}\sqrt{(r-r_+)(r-r_-)}d\theta, \tag{4.146}$$

$$\omega^1{}_3 = -\omega^3{}_1 = -\frac{r\sin\theta}{r^2+N^2}\sqrt{(r-r_+)(r-r_-)}d\phi, \tag{4.147}$$

$$\omega^2{}_3 = -\omega^3{}_2 = \frac{N(r-r_+)(r-r_-)}{(r^2+N^2)^2}dt + \left[\frac{2N^2(r-r_+)(r-r_-)}{(r^2+N^2)^2} - 1\right]\cos\theta d\phi. \tag{4.148}$$

As a consistency check, we can take the $N \to 0$ limit in these expressions, which yields the expected spin connection for the Schwarzschild solution. For the single copy solution of a magnetic monopole above, we considered a loop at constant time and radius in the equatorial plane $\theta = \pi/2$. The integral in the holonomy operator is then simply

$$\begin{aligned}\oint dx^\mu (\omega_\mu)_{ab} M^{ab} &= 2\oint d\phi[(\omega_\phi)_{02}M^{02} + (\omega_\phi)_{13}M^{13}] \\ &= 4\pi[(\omega_\phi)_{02}M^{02} + (\omega_\phi)_{13}M^{13}]. \end{aligned} \tag{4.149}$$

This yields a boost in the 0-2 plane and a rotation in the 1-3 plane, with coefficients $(\omega_\phi)_{02}$ and $(\omega_\phi)_{13}$ respectively, where

$$(\omega_\phi)_{02} = \frac{N}{r^2+N^2}\sqrt{(r-r_+)(r-r_-)}, \tag{4.150}$$

$$(\omega_\phi)_{13} = -\frac{r}{r^2+N^2}\sqrt{(r-r_+)(r-r_-)}. \tag{4.151}$$

The boost and rotation planes are thus mutually orthogonal. Such a transformation is conventionally known as a Lorentz four-screw. Our results here are in agreement with ref. [143], although we have adopted a different choice of vielbein.

The spin connection components of eqs. (4.150), (4.151) correspond to the full Taub-NUT solution. We wish to consider the pure NUT case, in which $M \to 0$, as this corresponds to the double copy of a magnetic charge. Taking this limit in eqs. (4.150), (4.151) yields

$$(\omega_\phi)_{02} = \frac{N}{r^2+N^2}\sqrt{r^2-N^2}, \tag{4.152}$$
$$(\omega_\phi)_{13} = -\frac{r}{r^2+N^2}\sqrt{r^2-N^2}, \tag{4.153}$$

such that the integral of eq. (4.149) reduces to

$$\oint dx^\mu (\omega_\mu)_{ab} M^{ab} = \frac{4\pi\sqrt{r^2-N^2}}{r^2+N^2}\left[N M^{02} - r M^{13}\right]. \tag{4.154}$$

Thus, the boost generator survives even in the case of a pure NUT charge. As explained in the previous section, the presence of the boost generator will potentially give rise to a SO(3, 1) holonomy group, unless the effects of the boost can be removed via a similarity transformation on the group elements. However, by considering other loops, boosts in different Cartesian directions are generated. To demonstrate this, note that the pure NUT metric has a Kerr-Schild form, and we can therefore make use of eq. (4.101) to express the integrand of the holonomy operator in terms of a Kerr-Schild graviton. The integral in eq. (4.154) is therefore equivalent to

$$-\oint dx^\mu \partial_\sigma (h_{\mu\nu}) M^{\nu\sigma} = -\oint dx^\mu \partial_\sigma(\phi k_\mu k_\nu) M^{\nu\sigma}. \tag{4.155}$$

Furthermore, for the pure NUT metric in Kerr-Schild form, we have that $k_0 = 1$, such that

$$\begin{aligned} -\oint dx^\mu \partial_\sigma(\phi k_\mu k_\nu) M^{\nu\sigma} &= -\oint dx^\mu \partial_\sigma(\phi k_\mu) M^{0\sigma} \\ &= -\oint dx^\mu \left[\partial_\sigma(\phi k_\mu) - \partial_\mu(\phi k_\sigma)\right] M^{0\sigma}, \end{aligned} \tag{4.156}$$

where in the second equality we have introduced a total derivative term that integrates to zero around a closed loop. We may now recognise the factor in square brackets as the abelian field strength associated with a gauge field that is the single copy of the Kerr-Schild graviton. Equation (4.156) can therefore be further rewritten as

$$\begin{aligned} -\oint dx^\mu \partial_\sigma(\phi k_\mu k_\nu) M^{\nu\sigma} &= -\oint dx^\mu F_{\sigma\mu} M^{\nu\sigma} \\ &= -\oint dx_j \epsilon_{ijk} B_k K_i \\ &= -\oint (d\boldsymbol{x} \times \boldsymbol{B}) \cdot \boldsymbol{K}. \end{aligned} \tag{4.157}$$

The physical content of this expression is made clear by considering a loop of fixed t and r, that is tilted relative to the equatorial plane. Such a loop is depicted in figure 4.2.

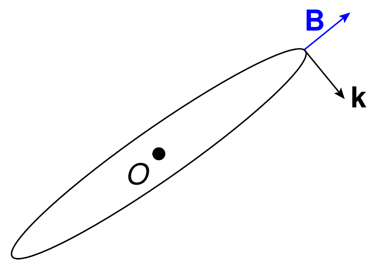


Figure 4.2: A closed spatial loop of fixed t and r, that is tilted with respect to the equatorial plane. O marks the origin. A monopole field $\boldsymbol{B}$ generates a boost in the direction $\boldsymbol{k}$.

The monopole $\boldsymbol{B}$ field is parallel to the radial direction, while $d\boldsymbol{x}$ is tangent to the integration contour. This generates a boost in the $\boldsymbol{k} \propto d\boldsymbol{x} \times \boldsymbol{B}$ direction. From the figure, it is clear that this is in the increasing θ direction. Therefore, boosts in the (t, θ) plane are generated for all such constant time loops, from which we can conclude that boosts in all Cartesian directions will be turned on in general. This gives rise to a SO(3, 1) holonomy group for the pure NUT solution, which does not match the SO(3) SCH group found for its single copy.

Naively, we might have hoped that these groups align with one another when the two solutions are related by the double copy. That this is not the case is a consequence of the physical properties of these two solutions. In ref. [143], the holonomy of Taub-NUT spacetime was studied using similar constant t loops. It was shown here that the generators M^{0i} arise due to the extrinsic curvature of space-like hypersurfaces within the spacetime. As spacelike hypersurfaces in the Taub-NUT solution have non-zero extrinsic curvature, both the "magnetic" M^{ij} and the "electric" M^{0i} generators are turned on in general. This is not the case for the magnetic monopole solution, which exists in gauge theory defined over Minkowski spacetime.

4.4.3 Self-dual solutions

We have now given two examples which probe the behaviour of the holonomy and SCH operators. In the case of the Schwarzschild / point charge system, both the holonomy and SCH groups remained in their maximal form of SO(3, 1). For the pure NUT / monopole system, the SCH group reduced while the holonomy group did not. In this section we demonstrate another possibility: that the SCH and holonomy groups reduce to mutually isomorphic subgroups of SO(3, 1). To illustrate this, we consider the case of self-dual solutions in gauge theory and gravity.

We begin with gravity. The Riemann tensor can be decomposed into its self-dual R^+ and anti-self dual R^- components via

$$R^{\pm}_{\mu\nu\rho\lambda} = (P^{\pm})^{\alpha\beta}{}_{\mu\nu} R_{\alpha\beta\rho\lambda}, \tag{4.158}$$

where we have defined the projectors

$$(P^{\pm})^{\alpha\beta}{}_{\mu\nu} = \frac{1}{2}\left(\delta^{\alpha}_{\mu}\,\delta^{\beta}_{\nu} - \delta^{\alpha}_{\nu}\,\delta^{\beta}_{\mu} \pm \sqrt{g}\epsilon^{\alpha\beta}{}_{\mu\nu}\right), \tag{4.159}$$

and g denotes the determinant of the metric. Note that in this section we work in Euclidean signature. We can use Stokes' theorem to rewrite the gravitational holonomy operator as

$$\begin{aligned}\Phi_{\omega} &= \mathcal{P}\exp\left(-\frac{i}{2}\iint_{\Sigma} d\Sigma^{\mu\nu}\, R_{cd\mu\nu}M^{cd}\right)\\ &= \mathcal{P}\exp\left(-\frac{i}{2}\iint_{\Sigma} d\Sigma^{\mu\nu}\,(R^{+}_{\rho\lambda\mu\nu} + R^{-}_{\rho\lambda\mu\nu})M^{\rho\lambda}\right),\end{aligned} \tag{4.160}$$

where $d\Sigma^{\mu\nu}$ is the area element of the area Σ bounded by the curve C. For a self-dual solution, the second term in this expression is zero by definition, such that the holonomy reduces to

$$\begin{aligned}\Phi_{\omega} &= \mathcal{P}\exp\left(-\frac{i}{2}\iint_{\Sigma} d\Sigma^{\mu\nu}\,(P^{+})^{\alpha\beta}{}_{\rho\lambda}R_{\alpha\beta\mu\nu}M^{\rho\lambda}\right)\\ &= \mathcal{P}\exp\left(-\frac{i}{2}\iint_{\Sigma} d\Sigma^{\mu\nu}\,R_{\alpha\beta\mu\nu}(M^{+})^{\alpha\beta}\right).\end{aligned} \tag{4.161}$$

Here we have defined two linearly independent sets of Lorentz generators, corresponding to the self-dual and anti-self-dual parts, via

$$(M^{\pm})^{\alpha\beta} = (P^{\pm})^{\alpha\beta}{}_{\rho\lambda}M^{\rho\lambda}. \tag{4.162}$$

This amounts to the Lie algebra isomorphism $\mathfrak{so}(4) \simeq \mathfrak{su}(2) \oplus \mathfrak{su}(2)$, where $(M^{\pm})^{\alpha\beta}$ are each generators of one of the $\mathfrak{su}(2)$ subalgebras. Thus, for self-dual solutions the holonomy group can immediately be seen to reduce to SU(2), and likewise for anti-self-dual solutions.

Self-dual gauge theory is the single copy of self-dual gravity [29]. The self-dual and anti-self-dual parts of the Yang-Mills field strength can be defined in an analogous manner to the Riemann tensor in eq. (4.158), via

$$F^{\pm}_{\mu\nu} = (P^{\pm})_{\mu\nu}{}^{\alpha\beta}F_{\alpha\beta}, \tag{4.163}$$

where the projectors are defined as in eq. (4.159), but with the general metric replaced with that of Euclidean flat space.[3] Self-dual solutions are those for which $F^{-}_{\mu\nu} = 0$, such that the SCH operator is

$$\begin{aligned}\exp\left[ig\oint_C ds F^{+}_{\mu\nu}M^{\mu\nu}\right] &= \exp\left[ig\oint_C ds (P^{+})_{\mu\nu}{}^{\alpha\beta}F_{\alpha\beta}M^{\mu\nu}\right]\\ &= \exp\left[ig\oint_C ds F_{\alpha\beta}(M^{+})^{\alpha\beta}\right]. \end{aligned} \tag{4.164}$$

Once again, only half of the Lorentz generators are turned on, so that the SCH group reduces to SU(2), as in gravity. The self-dual sector therefore provides an example where the SCH and holonomy groups both reduce to isomorphic subgroups.

4.4.4 Discussion

The identification of an operator corresponding to the single copy of the gravitational holonomy extends the catalogue of objects that are known to double copy, while also clarifying the question of what is to be considered as a gravitational Wilson line. However, from a practical perspective, this leaves a question: is the SCH operator useful for anything? In this section we have studied the nature of the holonomy and SCH groups for a collection of well-known exact double copies. We found three different outcomes. In two cases the holonomy and SCH group matched, taking their maximal form or mutually reducing, while in the case of the Taub-NUT metric and magnetic monopole solutions, we found that the SCH group reduces while the holonomy group does not. One might have hoped that the SCH operator could play a role in the identification of exact solutions for which a double copy is possible. However, the results of this section suggest that this may not be a fruitful line of thinking. Despite this fact, the SCH operator may still provide useful insights into the geometry of the double copy for exact solutions. As an operator it is valued in both spacetime and colour space, such that acts as a rotation of vectors both in the internal colour space and the tangent space of the spacetime manifold. In contrast to this, the gravitational holonomy acts as a rotation of vectors in the tangent space only. This is to be expected given that the information about the gauge group is removed when taking the double copy. However, a more formal study of what these replacements mean at the level of the fibre bundle geometry would be interesting, given that a geometrical understanding of the exact double copy (if any exists) is currently lacking.

[3] While the metric in eq. (4.159) now corresponds to flat Euclidean space, the $\sqrt{g}$ is still necessary as curvilinear coordinates may be employed.

Chapter 5

Non-perturbative insights from the self-dual sector

So far in this thesis, we have focused on the relationship between gauge theory and gravity that is revealed by the double copy. Broadly speaking, whether working at the level of scattering amplitudes or exact classical solutions, one moves from gauge theory to gravity by replacing colour with kinematic information. However, one can also move in the opposite direction. Starting with gauge theory and replacing kinematic with colour information yields biadjoint scalar theory, as depicted in figure 2.2. In all previous examples of the classical double copy, linearised solutions of this theory play a crucial role, appearing as the scalar functions that persist in the gauge and gravity solutions. However, for a fully non-perturbative understanding of the double copy, one would hope that it is be possible to identify non-linear solutions in all three theories. As the equations of motion for biadjoint scalar theory are simple in comparison to its double copy relatives, a possible approach to this question is to search directly for non-linear solutions to these equations. By assembling a catalogue of exact solutions in biadjoint scalar field theory, it may then be possible to identify counterparts in gauge and gravity theories, thereby shedding light on a non-linear exact double copy procedure.

This is a line of work initiated in ref. [144]. Here a set of exact biadjoint scalar solutions were found, the simplest of which corresponded to a spherically symmetric monopole-like object that is singular at the origin. Ref. [144] posited that this could by related to the non-abelian Wu-Yang monopole in gauge theory, however this was shown to not be the case in ref. [88] due to the fact that this gauge theory solution is related via a singular gauge transformation to a non-abelian dressing of the Dirac monopole, whose associated biadjoint field is already known. More exotic non-linear biadjoint solutions were found in refs. [145, 146], however the corresponding solutions in gauge theory or gravity, if they exist, are yet to be identified.

In this chapter we continue this line of work by searching for non-linear solutions to biadjoint scalar field theory in Euclidean spacetime. This is motivated by the fact that Euclidean solutions in gauge theory and gravity correspond to instantons. One might hope that if non-linear exact solutions exist in Euclidean biadjoint theory, it may be possible to relate them to known instanton solutions via the double copy. That this should be possible is further motivated by the fact that instantons correspond to (anti-)self-dual field configurations. The double copy is well understood in the self-dual sector, as reviewed in section 2.3.1, where perturbative classical solutions to self-dual Yang-Mills (SDYM) can be shown to double copy to those in self-dual gravity (SDG) [29].

We will begin by briefly reviewing some basic facts about instanton solutions in gauge theory and gravity, as well as the Kerr-Schild-like exact double copy between them. Following this, we will begin our search for non-linear solutions to Euclidean biadjoint scalar theory by looking for power-like spherically symmetric solutions, by analogy with the Lorentzian signature solutions of refs. [144, 145]. Here we find a general d-dimensional solution, however this has the curious property that it vanishes for $d = 4$. This turns out to be due to the fact that the power-like solutions already solve the *linearised* biadjoint field equations for $d = 4$. Furthermore, these linear solutions can be identified with a known double copy: they correspond to the zeroth copy of the Eguchi-Hanson instanton. In making this observation, we will be able to extend the double copy structure of these solutions, providing both a more general abelian-like single copy and a non-abelian embedding of this field. This non-abelian single copy maps to the same solution in gravity, mimicking the structure found for magnetic monopoles in chapter 3. Furthermore, our results will make contact with the two-dimensional non-perturbative double copy proposed in ref. [147]. From this, we will be able to interpret the sense in which the kinematic algebra is manifest in the context of certain exact four-dimensional solutions in SDYM and SDG.

5.1 Instantons in gauge theory and gravity

5.1.1 Exact solutions in self-dual gauge theory

We begin with an introduction to exact solutions to the self-dual Yang-Mills equations on Euclidean spacetime with finite action, which are known as instantons. There is a vast literature studying the properties and uses of gauge theory instantons (see e.g. refs. [92, 148, 149] for both classic and modern reviews). Here we give only a brief introduction to classical Euclidean gauge theory that will lay out the basic concepts necessary for later sections.

We work in four-dimensional Euclidean spacetime, where the metric is simply $\delta_{\mu\nu}$. As is common in the instanton literature, when in Euclidean spacetime we let Greek indices take the values $\mu, \nu, ... = 1, 2, 3, 4$, such that a general vector is

$$x^\mu = (x, y, z, \tau), \tag{5.1}$$

where τ is the Euclidean time coordinate, $x^4 = ix^0 \equiv \tau$. As we are in Euclidean signature there is now no distinction between upper and lower indices.

Consider pure $SU(N)$ Yang-Mills theory with the action

$$S = -\frac{1}{4}\int d^4x\,\mathrm{tr}(F_{\mu\nu}F_{\mu\nu}). \tag{5.2}$$

Variation of the action yields the Euclidean Yang-Mills field equations

$$D_\mu F_{\mu\nu} = 0, \tag{5.3}$$

where D_μ is the gauge covariant derivative. Instantons are exact finite action solutions to these equations of motion. The stipulation of a finite action fixes the asymptotic behaviour of the gauge field solutions to eq. (5.3). As the integrand of the action in eq. (5.2) goes like F^2, we require that the field strength must decrease at infinity faster than x^{-2}. For this to be the case, the gauge field must be asymptotically pure gauge, such that as $x \to \infty$

$$A_\mu \to iU\partial_\mu U^{-1}. \tag{5.4}$$

where $U \in \mathrm{SU}(N)$. Finite action solutions are therefore asymptotically classified by elements of the gauge group, which act as maps from the spacetime boundary to the gauge group

$$U : S^3 \to \mathrm{SU}(N). \tag{5.5}$$

Such maps are classified by the third homotopy group π_3, which for $\mathrm{SU}(N)$ is

$$\pi_3(\mathrm{SU}(N)) = \mathbb{Z}. \tag{5.6}$$

An integer $k \in \mathbb{Z}$ thus classifies the solutions, and is given by the integral

$$k = \frac{1}{16\pi^2} \int d^4x \, \mathrm{tr}(F_{\mu\nu}\tilde{F}_{\mu\nu}), \tag{5.7}$$

where $\tilde{F}$ is the dual field strength

$$\tilde{F} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F_{\rho\sigma}. \tag{5.8}$$

The integer k goes by many names: the winding number, instanton number, topological charge, or Pontryagin class.

We have seen that instanton solutions are characterised by an integer value k. How do we actually find such solutions? We could search for them by attempting to solve eq. (5.3). This is a second-order equation however, and things will be greatly expedited by reformulating eq. (5.3) as a first-order equation. To do this, we note that due to the property

$$\mathrm{tr}(F^2) = \mathrm{tr}(\tilde{F}^2), \tag{5.9}$$

we can rewrite the Euclidean action of eq. (5.2) as

$$\begin{aligned} S &= \frac{1}{8}\int d^4x \, \mathrm{tr}(F^2 + \tilde{F}^2) \\ &= \frac{1}{8}\int d^4x \left[\mathrm{tr}(F_{\mu\nu} \mp \tilde{F}_{\mu\nu})^2 \pm 2\mathrm{tr}(F_{\mu\nu}\tilde{F}_{\mu\nu})\right] \\ &= \pm 4\pi^2 k + \frac{1}{8}\int d^4x \, \mathrm{tr}(F_{\mu\nu} \mp \tilde{F}_{\mu\nu})^2, \end{aligned} \tag{5.10}$$

where in the final equality we have identified the instanton number k from eq. (5.7). By noting that the second term in the final line of this expression is a total square, we can place a lower bound on the action, such that

$$S \geq 4\pi^2|k|. \tag{5.11}$$

From eq. (5.10) we see that this bound is saturated when

$$F_{\mu\nu} \mp \tilde{F}_{\mu\nu} = 0. \tag{5.12}$$

Thus, for a given topological sector defined by k, the action is minimised by solving the following expressions

$$F_{\mu\nu} = \tilde{F}_{\mu\nu}, \quad k > 0 \tag{5.13}$$

$$-F_{\mu\nu} = \tilde{F}_{\mu\nu}, \quad k < 0. \tag{5.14}$$

Conventionally, the $k > 0$ solutions are referred to as instantons while $k < 0$ solutions are referred to as anti-instantons. Eqs. (5.13) and (5.14) are known as the self-duality and anti-self duality conditions respectively. Solutions to these expressions necessarily solve the Euclidean Yang-Mills equations of eq. (5.3).

5.1.2 Exact solutions in self-dual gravity

The definition of instanton solutions in gravitational theories is less rigorous than that in gauge theory. A gravitational instanton is always a solution to the Euclidean Einstein equations, however the additional properties vary from case to case. For example, solutions with either finite or non-finite action, self-dual or non-self-dual curvature, and asymptotically locally Euclidean (ALE) or non-ALE behaviour are all often referred to as gravitational instantons. For our purposes, in analogy to Yang-Mills instantons, we consider gravitational instantons to be four-dimensional Riemannian metrics which satisfy the Euclidean Einstein equations and the (anti-)self-duality conditions

$$R_{\mu\nu\rho\sigma} = \pm\tilde{R}_{\mu\nu\rho\sigma}, \tag{5.15}$$

where $R_{\mu\nu\rho\sigma}$ are the components of the Riemann curvature tensor. The dual curvature tensor is defined as

$$\tilde{R}_{\mu\nu\rho\sigma} = \frac{1}{2}\epsilon_{\mu\nu\gamma\delta}R_{\gamma\delta\rho\sigma}. \tag{5.16}$$

One might wonder about the origin of this definition, seeing as the operation of the Hodge dual maps a two-form to another two-form. Indeed, eq. (5.15) arises from imposing the (anti-)self-duality conditions on the curvature two-form

$$R_{ab} = \pm\tilde{R}_{ab}, \tag{5.17}$$

which furthermore corresponds to an (anti-)self-dual spin connection. The (anti-)self-duality of the curvature along with the Bianchi identity implies that the corresponding metrics are solutions to the vacuum Einstein equations.

In further analogy to Yang-Mills instantons, gravitational instanton solutions fall into distinct topological sectors. These are classified by two topological invariants, the

Pontryagin number and the Euler characteristic, given respectively by [150]

$$p = \frac{1}{32\pi^2} \int d^4x \sqrt{g}\, \epsilon_{\mu\nu\alpha\beta} R_{\mu\nu\rho\sigma} R_{\rho\sigma\alpha\beta}, \tag{5.18}$$

$$\chi = \frac{1}{128\pi^2} \int d^4x \sqrt{g}\, \epsilon_{\mu\nu\alpha\beta} \epsilon_{\rho\sigma\gamma\delta} R_{\mu\nu\rho\sigma} R_{\alpha\beta\gamma\delta}. \tag{5.19}$$

A variety of gravitational instanton solutions are known (see the seminal works of refs. [150–154], and ref. [97] for a modern review).

5.1.3 The double copy for exact self-dual solutions

In ref. [21] a Kerr-Schild-like approach to double-copying exact self-dual solutions was proposed. In this approach, the Kerr-Schild graviton is formulated in terms of a differential operator $\hat{k}_\mu$, such that the full metric takes the form

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa \hat{k}_\mu \hat{k}_\nu \phi. \tag{5.20}$$

In the standard Kerr-Schild double copy, the graviton decomposes into a local product in position space, whereas here the use of differential operators corresponds to a local product in momentum space.[1] The operator $\hat{k}_\mu$ is taken to satisfy the following properties

$$\hat{k}^2 = 0, \qquad \partial \cdot \hat{k} = 0, \qquad [\hat{k}_\mu, \hat{k}_\nu] = 0. \tag{5.21}$$

We now go to the light-cone coordinates

$$u = t - z, \quad v = t + z, \quad w = x + iy, \quad \bar{w} = x - iy. \tag{5.22}$$

Note that for this section we are working in Minkowski spacetime, as opposed to Euclidean, so as to match the literature. One can show that for a metric satisfying the self-duality condition of eq. (5.15), the differential operator takes the form

$$\hat{k}_u = 0, \quad \hat{k}_v = \frac{1}{4}\partial_w, \quad \hat{k}_w = 0, \quad \hat{k}_{\bar{w}} = \frac{1}{4}\partial_u, \tag{5.23}$$

where the numerical coefficients are chosen. Furthermore, the Einstein equations reduce to

$$\partial^2\phi - \frac{\kappa}{2}(\hat{k}^\mu \hat{k}^\nu \phi)(\partial_\mu \partial_\nu \phi) = 0, \tag{5.24}$$

which, for the operator in eq. (5.23) reduces to the Plebanski equation for SDG of eq. (2.47).

[1] See refs. [82, 155] for recent discussions about when and why the classical double copy takes a local position-space form.

One can single copy the self-dual graviton in eq. (5.20) by following an analogous approach to the Kerr-Schild single copy. That is, we consider a gauge field

$$A^a_\mu = \hat{k}_\mu \phi^a, \tag{5.25}$$

where the scalar is valued in the Lie algebra. Substituting this ansatz into the Yang-Mills equations yields

$$\hat{k}_\nu \partial^2 \phi^a + 2g f^{abc}(\hat{k}^\mu \phi^b)(\hat{k}_\nu \partial_\mu \phi^c) = 0, \tag{5.26}$$

which is equivalent to the SDYM equations of eq. (2.46) for the operator in eq. (5.23). We therefore see that exact solutions to self-dual gauge theory and gravity can be double copied in a Kerr-Schild-like manner. This aligns with the results of ref. [29], which we reviewed in section 2.3.1, in which perturbative solutions to the classical equations of motion in the self-dual sectors where shown exhibit manifest colour-kinematics duality.

5.2 Exact solutions in Euclidean biadjoint scalar theory

Biadjoint scalar field theory was introduced in section 2.3.2. It is described by a cubic Lagrangian, which gives rise to the Lorentzian signature equations of motion in eq. (2.75). The biadjoint scalar theory is an unphysical theory, with energies that are unbounded from below such that exact solutions are dynamically unstable. However, as we have previously discussed, the theory is intimately tied to the double copy in that it is obtained from Yang-Mills through the replacement of kinematic with colour information. That is, it corresponds to the zeroth copy of Yang-Mills theory. When considering exact classical solutions, the biadjoint scalar theory still arises as the zeroth copy of well-behaved, physical solutions in gauge theory. Thus, in probing the non-perturbative structure of the double copy, it is useful to catalogue exact solutions in biadjoint scalar field theory, regardless of their non-physical properties. This line of work was initiated in ref. [144] and further developed in refs. [3, 88, 145, 146]. In all of these works, biadjoint scalar theory was considered in Lorentzian signature. Here we look for Euclidean solutions.

We begin by transforming the Lorentzian biadjoint scalar theory equations of motion in eq. (2.75) to Euclidean signature. This is done via the Wick rotation $t \to i\tau$, which amounts to

$$\partial^2 \to \Delta, \tag{5.27}$$

where Δ is the Laplacian operator. For now, we work in d-dimensions. The Euclidean equations of motion are then

$$\Delta\Phi^{aa'} + y f^{abc}\tilde{f}^{a'b'c'}\Phi^{bb'}\Phi^{cc'} = 0. \tag{5.28}$$

We search for exact solutions to this equation following the approach suggested in ref. [144]. Solutions to biadjoint scalar theory are valued in two, possibly distinct gauge groups, G and $\tilde{G}$, each with their own sets of structure constants. Let us first assume that these groups coincide, such that

$$f^{abc} = \tilde{f}^{abc}. \tag{5.29}$$

The structure constants satisfy the identity

$$f^{abc}f^{a'bc} = T_A\delta^{aa'}, \tag{5.30}$$

where T_A is a constant dependent on the common gauge group G and the normalisation of the generators. Here, we will restrict to spherically symmetric solutions, via the ansatz

$$\Phi^{aa'} = \frac{\delta^{aa'}}{yT_A}f(r), \quad r^2 = x_\mu x^\mu. \tag{5.31}$$

Substituting this into the equations of motion in eq. (5.28) yields

$$\frac{1}{r^{d-1}}\frac{d}{dr}\left(r^{d-1}\frac{df(r)}{dr}\right) + f^2(r) = 0, \tag{5.32}$$

where the d-dimensional Laplacian in spherical coordinates has been used:

$$\Delta f = \frac{1}{r^{d-1}}\frac{\partial}{\partial r}\left(r^{d-1}\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\Delta_{S^{d-1}}f. \tag{5.33}$$

The second term contains the Laplace-Beltrami operator $\Delta_{S^{d-1}}$, which depends on angular coordinates only. As we are considering spherically symmetric solutions this term does not contribute in eq. (5.32).

Let us now restrict to power-like solutions, via the radial function

$$f(r) = Ar^\alpha, \tag{5.34}$$

where A and α are constants. Substituting this into eq. (5.32) yields

$$A\alpha(d+\alpha-2)r^{\alpha-2} + A^2r^{2\alpha} = 0. \tag{5.35}$$

Requiring that this holds for all $r > 0$ enforces

$$\alpha = -2 \quad \Rightarrow \quad A[A - 2(d-4)] = 0, \tag{5.36}$$

which yields two solutions

$$A = 0, \quad A = 2(d-4). \tag{5.37}$$

Thus, with eqs. (5.31) and (5.34), we have two solutions to the Euclidean equations of motion. The first, corresponding to $A = 0$, is the trivial vacuum solution $\Phi^{aa'} = 0$. The second is a non-trivial power-like solution

$$\Phi^{aa'} = \frac{2\delta^{aa'}}{yT_A}\frac{d-4}{r^2}. \tag{5.38}$$

Note that this is a d-dimensional solution, and thus the radial dependence of r^{-2} is common to all spacetime dimensions. This fact can be confirmed by dimensional analysis. As d varies for solutions with an inverse power of the coupling, the dimension of the fields and coupling vary so as to fix the radial power. This is in contrast to the linearised field equations, whose solutions must depend on a radial power that varies as the dimension changes, so as to maintain the correct dimensions for the field.

Let us perform a brief consistency check of the solution in eq. (5.38) by comparing it to the results of ref. [144]. In this paper static monopole-like solutions to biadjoint scalar field theory in Lorentzian signature were found in four-dimensions. As these solutions are static, the Lorentzian equation of motion in eq. (2.75) for $d = 4$ reduces to the Euclidean equation in eq. (5.28) for $d = 3$. Taking $d = 3$ in eq. (5.38) yields

$$\Phi^{aa'}\Big|_{d=3} = -\frac{2\delta^{aa'}}{yT_A r^2}. \tag{5.39}$$

This is precisely eq. (11) in ref. [144].

A peculiar feature of our solution in eq. (5.38) is the presence of $(d-4)$ in the numerator. We are most interested in $d = 4$ corresponding to four-dimensional Euclidean spacetime, but this factor tells us that there are no such non-trivial power-like solutions to eq. (5.28). To investigate this further, let us take influence from the Lorentzian signature work of ref. [145], and generalise our spherically symmetric ansatz. In this paper, the following ansatz for the radial function was considered

$$f(r) = \frac{K(r) - 1}{r^2}, \tag{5.40}$$

where $K(r)$ is finite for all r. This ansatz partially screens the divergence at the origin, and aligns with the trivial solution for $K(r) \to 1$. Note that finite energy static solutions in biadjoint scalar theory are prohibited by Derrick's theorem [156]. Taking this form for $f(r)$ in eq. (5.33) yields

$$r^2K''(r) + (d-5)rK' + K^2 - 2(d-3)K + 2d - 7 = 0, \tag{5.41}$$

where the prime denotes differentiation with respect to r. Let us now introduce a variable ξ, such that

$$r = e^{-\xi}, \quad -\infty < \xi < \infty, \tag{5.42}$$

which reduces eq. (5.41) to

$$\frac{\partial^2 K}{\partial \xi^2} - (d-6)\frac{\partial K}{\partial \xi} + (K-1)(K-2d+7) = 0. \tag{5.43}$$

This is a non-linear second-order differential equation. In general it is not analytically solvable. Ref. [145] showed that via a further transformation it can be recognised as an Abel equation of the second kind, although not one that has an analytic solution in terms of known functions. We can, however, gain valuable information from eq. (5.43) by visualising its solutions via a method similar to that used in ref. [157] in the construction of exact solutions to pure Yang-Mills. The approach is to reformulate eq. (5.43) as two coupled first-order differential equations. We introduce a new parameter ψ via

$$\psi = \frac{\partial K}{\partial \xi}. \tag{5.44}$$

Equation (5.43) then corresponds to the coupled equations

$$\left(\frac{\partial K}{\partial \xi}, \frac{\partial \psi}{\partial \xi}\right) = \Big(\psi, (d-6)\psi - (K-1)(K-2d+7)\Big). \tag{5.45}$$

This defines a vector field in the (K, ψ) plane, whose integral curves correspond to solutions of eq. (5.43). In figure 5.1, we show plots of these curves for $d = 2, 3, 4, 5$. For general d, the vector field of eq. (5.45) vanishes at two fixed points:

$$(K, \psi) = (1, 0), \tag{5.46}$$

$$(K, \psi) = (2d-7, 0). \tag{5.47}$$

These correspond to the trivial solution and the non-trivial power-like solution in eq. (5.38). From the plots it can be seen that as d increases the non-trivial solution moves to right along the $\psi = 0$ axis in the (K, ψ) plane. However, for $d = 4$ the two fixed points coincide, so that there is only the trivial solution.

Solutions K that are finite for all r correspond to integral curves that are bounded in the (K, ψ) plane, i.e. those that interpolate between the fixed points. For $d \neq 4$ there is always a single bounded curve, connecting the points $(1, 0)$ and $(2d - 7, 0)$. This broadens the implications of the absence of two fixed points for $d = 4$: not only are there no non-trivial spherically symmetric power-like solutions, but there are also no non-trivial extended solutions of the form in eq. (5.40).

5.3 Revisiting the single copy of the Eguchi-Hanson instanton

We have seen in the previous section that there are no non-trivial power-like or spherically symmetric solutions to four-dimensional Euclidean biadjoint scalar theory. There turns out to be a very simple reason why this is the case. Recall that in eq. (5.38), we saw that spherically symmetric power-like solutions to the full non-linear equations of motion always have a radial dependence that goes like $\sim r^{-2}$, regardless of the spacetime dimension. However, for $d = 4$, r^{-2} is a harmonic function. Thus, by definition, the Laplacian of solutions with this radial dependence vanish for $r \neq 0$. To make this explicit, recall the form of the Laplacian in spherical coordinates in eq. (5.33). This implies that in d dimensions

$$\Delta r^{-n} = n(n + 2 - d) r^{-n-2}, \tag{5.48}$$

which vanishes for $r > 0$, $n = 2$, and $d = 4$. Fields $\Phi^{aa'}$ with r^{-2} dependence in four dimensions therefore solve the *linearised* biadjoint equations

$$\Delta \Phi^{aa'} = 0. \tag{5.49}$$

Thus, there are no non-trivial power-like solutions to the full non-linear equations of motion in four dimensions. While this argument circumvents the need for the analysis in the previous section, we believe that it was still worthwhile. There are non-trivial power-like solutions in other numbers of dimensions, which deserve further study. Furthermore, the fact that bounded extended solutions in $d = 4$ are also absent would have been missed by this simple argument.

Solutions to the linearised biadjoint equations of motion can be promoted to solutions to the full non-linear equations by dressing them with constant colour vectors $\{c^a, \tilde{c}^{a'}\}$ [21]. Indeed this is exactly the form of the solutions obtained via the Kerr-Schild double copy.

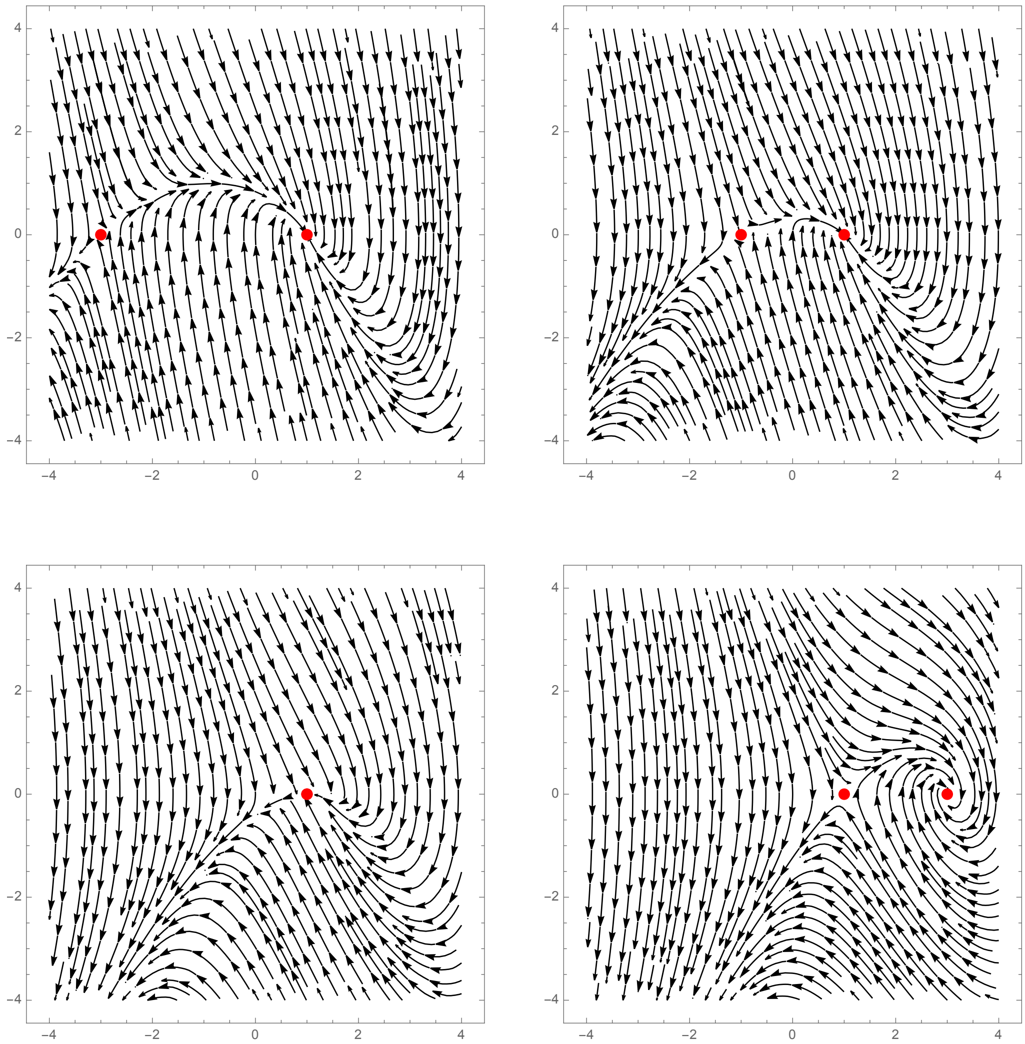


Figure 5.1: Plots of the vector field of eq. (5.45) in the (K, ψ) plane for $d = 2, 3, 4, 5$ respectively. Integral curves correspond to solutions of the differential equation in eq. (5.43). The red dots correspond to the fixed point solutions $(K, \psi) = (1, 0)$ and $(K, \psi) = (2d - 7, 0)$.

For a general harmonic solution, we therefore have

$$\Phi^{aa'} = \frac{\alpha c^a \tilde{c}^{a'}}{r^2}, \tag{5.50}$$

where α is an arbitrary constant. As expected, in the substitution of this ansatz into the equations of motion in eq. (2.75), the non-linear term vanishes and we are left with a solution to the linearised equations. Due to the precedence of this form of solutions in the classical double copy, it is natural to ask whether there are gauge and gravity solutions for which eq. (5.50) constitutes the zeroth copy. In fact, the solution is already known: it is related to the Eguchi-Hanson instanton, a known gravitational instanton in general relativity. The Eguchi-Hanson metric takes a particularly nice form in the following coordinate system [22, 74]

$$u = \frac{\tau - iz}{\sqrt{2}}, \quad v = \frac{\tau + iz}{\sqrt{2}}, \quad X = \frac{ix - y}{\sqrt{2}}, \quad Y = \frac{ix + y}{\sqrt{2}}. \tag{5.51}$$

The metric then takes a Kerr-Schild form in which the graviton can be written as [22]

$$h_{\mu\nu} = \phi k_\mu k_\nu, \quad \phi = \frac{\lambda}{(uv - XY)}, \quad k_\mu = \frac{1}{(uv - XY)}(v, 0, 0, -X). \tag{5.52}$$

The Kerr-Schild vector k_μ satisfies the usual null and geodesic conditions. The single and zeroth copies are then straightforwardly obtained following the standard procedure [21], giving rise to the following gauge and biadjoint scalar fields

$$A^a_\mu = c^a \phi k_\mu, \quad \Phi^{aa'} = c^a \tilde{c}^{a'} \phi. \tag{5.53}$$

where c^a and $\tilde{c}^{a'}$ are constant colour vectors. As is standard for exact classical double copies, these solutions linearise their respective equations of motion. Translating the scalar field in eq. (5.52) into Euclidean Cartesian coordinates, we find

$$\phi = \frac{2\lambda}{r^2}. \tag{5.54}$$

Thus, the zeroth copy of the Eguchi-Hanson metric in eq. (5.53) is equal to the power-like solution in eq. (5.50) with $\alpha = 2\lambda$.

For our purposes, it will be useful to reformulate this result in an alternative way. First, let us rewrite the single copy gauge field in eq. (5.53) in terms of an abelian gauge field A_μ:

$$A^a_\mu = c^a A_\mu, \quad A_\mu = \frac{\lambda}{(uv - XY)^2}\,(v, 0, 0, -X)\,. \tag{5.55}$$

We may then identify a differential operator $\hat{k}_\mu$, such that

$$A_\mu = \hat{k}_\mu \phi, \quad \hat{k}_\mu = -(\partial_u, 0, 0, \partial_Y), \tag{5.56}$$

where the single copy is written in terms of the operators action on the scalar field ϕ. Furthermore, the Eguchi-Hanson graviton can be written in terms of the same operator via

$$h_{\mu\nu} = \hat{k}_\mu \hat{k}_\nu \phi. \tag{5.57}$$

This construction appears to mirror the Kerr-Schild-like approach to the double copy for exact self-dual solutions reviewed in section 5.1.3, in which differential operators play a role analogous to the Kerr-Schild vectors. Indeed, such a formulation is possible given that the Eguchi-Hanson metric is an example of a self-dual solution [74]. Recall, however, that the differential operator used in this construction was required to satisfy the properties in eq. (5.21). The operator used here indeed meets the first criterion, such that $\hat{k}^2 = 0$. However, the second is not satisfied, and we find that

$$\partial \cdot \hat{k} = -\frac{1}{2}\Delta. \tag{5.58}$$

It is therefore unclear whether the gauge and graviton fields of eqs. (5.55), (5.56), and (5.57) constitute a special case of the self-dual double copy proposed in ref. [21].

To clarify this apparent mismatch, it will be useful to transform the differential operator of eq. (5.56) into Cartesian coordinates, where it takes the form

$$\hat{k}_\mu = -\frac{1}{2}\left(\partial_x + i\partial_y, \partial_y - i\partial_x, \partial_z - i\partial_\tau, \partial_\tau + i\partial_z\right). \tag{5.59}$$

We now notice that this can be written in the compact form

$$\hat{k}_\mu = -\frac{1}{2}\left(\delta_{\mu\nu} + i\bar{\eta}^3_{\mu\nu}\right)\partial_\nu, \tag{5.60}$$

where $\bar{\eta}^3_{\mu\nu}$ is an example of a 't Hooft symbol $\{\bar{\eta}^a_{\mu\nu}\}$, given explicitly by

$$\bar{\eta}^a_{\mu\nu} = \epsilon^a{}_{\mu\nu 4} - \delta^a_\mu \delta_{\nu 4} + \delta^a_\nu \delta_{\mu 4} \tag{5.61}$$

't Hooft symbols are common in the study of gauge theory instantons and will play an important role in the following sections. We will therefore review their properties in more detail in the next section. For now, it is enough to know that they form a representation of an SU(2) subalgebra of SO(4), the Lorentz group in Euclidean signature. Thus, in eq. (5.60), the 't Hooft symbol enacts a particular "rotation" on the derivative operator ∂_μ. Let us then consider a particular rotation on $\hat{k}_\mu$ itself, to obtain a second operator $\hat{k}'_\mu$:

$$\begin{aligned}\hat{k}'_\mu &= -\bar{\eta}^2_{\mu\nu}\hat{k}_\nu \\ &= \frac{1}{2}\left(\bar{\eta}^2_{\mu\nu} - i\bar{\eta}^1_{\mu\nu}\right)\partial_\nu,\end{aligned} \tag{5.62}$$

which in the coordinates of eq. (5.51) is given by

$$\hat{k}'_\mu = (0, \partial_Y, \partial_u, 0). \tag{5.63}$$

With this new operator, it is straightforward to confirm that

$$\hat{k}'^2 = 0, \quad \partial \cdot \hat{k}' = 0, \tag{5.64}$$

such that the conditions in eq. (5.21) are satisfied. Furthermore, $\hat{k}'_\mu$ is simply a coordinate transformation of $\hat{k}_\mu$ to a frame whose coordinates are

$$x'_\mu = -\bar{\eta}^2_{\mu\nu}x_\nu \implies (x', y', z', \tau') = (z, \tau, -x, -y). \tag{5.65}$$

We then have

$$(r')^2 = x'_\mu\, x'_\mu = x_\mu\, x_\mu = r^2, \tag{5.66}$$

such that the Eguchi-Hanson graviton and its single copy in the primed coordinate system take the form

$$A'_\mu = c^a\hat{k}'_\mu\phi(r'), \quad h'_{\mu\nu} = \hat{k}'_\mu\hat{k}'_\nu\phi(r'). \tag{5.67}$$

This is now a special case of the exact self-dual double copy reviewed in section 5.1.3.

From the discussion in this section, we tied the absence of non-linear power-like solutions in Euclidean biadjoint scalar theory to the double copy, by noting that simple power-like solutions already correspond to a known zeroth copy, that of the Eguchi-Hanson instanton, for which the biadjoint theory equations of motion linearise. As the Eguchi-Hanson metric is a self-dual solution, its single and zeroth copies are naturally described in terms of differential operators satisfying eqs. (5.21), such that it corresponds to a special case of the exact self-dual double copy construction proposed in ref. [21]. Particularly compelling for our present purposes is the appearance of 't Hooft symbols in the differential operators from which the fields related by the double copy are built. We will find that their presence allows for the construction of a general ansatz for double copying certain self-dual solutions, as well as a straightforward generalisation to non-abelian single copies. However, due to the importance of the 't Hooft symbols in this story, we will now briefly review their properties and provide some useful identities.

5.4 ’t Hooft symbols

The ’t Hooft symbols, first introduced in ref. [158], appear frequently in the study of instanton solutions in gauge theory. In Euclidean spacetime, the Lorentz group is SO(4). This is a six-dimensional group, with three rotations $\{J_i\}$ in the (x_i, x_j) plane, and three rotations $\{K_i\}$ in the (x_i, x_4) plane (recall that $x^4 = \tau$ is our Euclidean time coordinate). Note that the latter are the analogue of boosts in Lorentzian signature. We can form linear combinations of these two sets of rotations

$$M_i = \frac{1}{2}(J_i + K_i), \quad N_i = \frac{1}{2}(J_i - K_i), \tag{5.68}$$

which furnish two independent SU(2) subalgebras, such that

$$[M_i, M_j] = -\epsilon_{ijk} M_k, \quad [N_i, N_j] = -\epsilon_{ijk} N_k, \quad [M_i, N_j] = 0. \tag{5.69}$$

The ’t Hooft symbols form a representation of the generators $\{M_i, N_i\}$ acting on four-dimensional vectors. The are given explicitly by

$$\eta^a_{\mu\nu} = \epsilon^a{}_{\mu\nu 4} + \delta^a_\mu \delta_{\nu 4} - \delta^a_\nu \delta_{\mu 4}, \tag{5.70}$$

$$\bar{\eta}^a_{\mu\nu} = \epsilon^a{}_{\mu\nu 4} - \delta^a_\mu \delta_{\nu 4} + \delta^a_\nu \delta_{\mu 4}, \tag{5.71}$$

and satisfy

$$[\eta^a, \eta^b] = -2\epsilon^{abc}\eta^c, \quad [\bar{\eta}^a, \bar{\eta}^b] = -2\epsilon^{abc}\bar{\eta}^c, \quad [\eta^a, \bar{\eta}^b] = 0, \tag{5.72}$$

where the indices $a, b, c, ... \in \{1, 2, 3\}$ and we have here used matrix notation. Thus, we can identify

$$M_i = \frac{1}{2}\bar{\eta}^i_{\mu\nu}, \quad N_i = \frac{1}{2}\eta^i_{\mu\nu}. \tag{5.73}$$

Note that $\eta^a_{\mu\nu}$ and $\bar{\eta}^a_{\mu\nu}$ are antisymmetric

$$\eta^a_{\mu\nu} = -\eta^a_{\nu\mu}, \qquad \bar{\eta}^a_{\mu\nu} = -\bar{\eta}^a_{\nu\mu} \tag{5.74}$$

and self-dual and anti-self-dual respectively:

$$\eta^a_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\eta^a_{\rho\sigma}, \qquad \bar{\eta}^a_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\bar{\eta}^a_{\rho\sigma}. \tag{5.75}$$

Furthermore, from eqs. (5.70) and (5.71), we find their explicit matrix form to be

$$\eta^1_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \eta^2_{\mu\nu} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \eta^3_{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix};$$

$$\bar{\eta}^1_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \bar{\eta}^2_{\mu\nu} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \bar{\eta}^3_{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{5.76}$$

The 't Hooft symbols satisfy a number of useful properties. In particular, the following identities will be used frequently in the following sections:

$$\begin{aligned} \bar{\eta}^a_{\mu\nu}\,\bar{\eta}^a_{\rho\sigma} &= \delta_{\mu\rho}\,\delta_{\nu\sigma} - \delta_{\mu\sigma}\,\delta_{\nu\rho} - \epsilon_{\mu\nu\rho\sigma}; \\ \bar{\eta}^a_{\mu\rho}\,\bar{\eta}^b_{\mu\sigma} &= \delta^{ab}\,\delta_{\rho\sigma} + \epsilon^{abc}\bar{\eta}^c_{\rho\sigma}; \\ \epsilon^{abc}\,\bar{\eta}^b_{\mu\nu}\,\bar{\eta}^c_{\rho\sigma} &= \delta_{\mu\rho}\bar{\eta}^a_{\nu\sigma} + \delta_{\nu\sigma}\bar{\eta}^a_{\mu\rho} - \delta_{\mu\sigma}\bar{\eta}^a_{\nu\rho} - \delta_{\nu\rho}\bar{\eta}^a_{\mu\sigma} \\ \bar{\eta}^a_{\mu\nu}\bar{\eta}^b_{\mu\nu} &= 4\delta^{ab}. \end{aligned} \tag{5.77}$$

Further identities may be found in e.g. appendix B of ref. [159].

5.5 A general ansatz for double-copying self-dual solutions

In pure Yang-Mills there are many instanton solutions, the full classification of which is given by the ADHM construction [160]. For our purposes, we will consider a large family of SU(2) instantons that are constructed according to the so-called 't Hooft ansatz [158]. This involves dressing a vector field V_μ with the 't Hooft symbols, such that

$$A^a_\mu = -\bar{\eta}^a_{\mu\nu}V_\nu, \quad A^a_\mu = -\eta^a_{\mu\nu}V_\nu, \tag{5.78}$$

are self-dual and anti-self-dual fields respectively. When the 't Hooft symbols appeared previously in the definition of the differential operator in eq. (5.60), they were acting as a representation of a particular spacetime rotation, with the upper index labelling the infinitesimal rotation in question. Here, however, we can interpret the upper indices as adjoint indices associated with the SU(2) gauge group. This is possible due to the fact that $\eta^a_{\mu\nu}$ and $\bar{\eta}^a_{\mu\nu}$ each furnish a representation of SU(2), as outlined in the previous section. For ease of reference, we will consider only the self-dual case. Substituting the self-dual field in eq. (5.78) into the Yang-Mills equations yields the following constraints on the field V_μ:

$$\partial_\mu V_\mu + V_\mu V_\mu = 0, \tag{5.79}$$

$$f_{\mu\nu} - \tilde{f}_{\mu\nu} = 0, \tag{5.80}$$

where we have defined

$$f_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu. \tag{5.81}$$

We can recognise the quantity $f_{\mu\nu}$ as being analogous to an abelian-like field strength. However, it is not immediately possible to identify V_μ as an abelian gauge field, given that there is no guarantee in general that $f_{\mu\nu}$ satisfies the Maxwell equation,

$$\partial_\mu f_{\mu\nu} = 0. \tag{5.82}$$

In the standard approach, eq. (5.79) is satisfied by taking V_μ to be the gradient of the logarithm of a harmonic scalar:

$$V_\mu = \partial_\mu \log V, \quad \Delta V = 0. \tag{5.83}$$

However, following our discussion of the Eguchi-Hanson instanton, we can propose an alternative ansatz. Consider constructing a vector field according to the prescription:

$$A_\mu = \hat{k}_\mu \phi, \quad \hat{k}_\mu = \left(A\delta_{\mu\nu} + B_i\,\bar{\eta}^i_{\mu\nu}\right)\partial_\nu, \quad i \in \{1,2,3\}, \tag{5.84}$$

where the scalar A and vector B are possibly complex. This generalises the definitions of $\hat{k}_\mu$ and $\hat{k}'_\mu$ given previously, such that eqs. (5.60) and (5.62) are special cases of eq. (5.84). In terms of this ansatz, the quantity $f_{\mu\nu}$ is

$$f_{\mu\nu} = B_i\left(\bar{\eta}^i_{\nu\rho}\partial_\mu - \bar{\eta}^i_{\mu\rho}\partial_\nu\right)\partial_\rho\phi, \tag{5.85}$$

from which we find

$$\tilde{f}_{\mu\nu} = f_{\mu\nu} + B_i\bar{\eta}^i_{\mu\nu}\Delta\phi. \tag{5.86}$$

Thus, for A_μ constructed as in eq. (5.84), the constraint in eq. (5.80) is satisfied when ϕ is a harmonic function

$$\Delta\phi = 0 \iff f_{\mu\nu} = \tilde{f}_{\mu\nu}. \tag{5.87}$$

The requirement that ϕ is harmonic also means that eq. (5.84) solves the Maxwell equation

$$\partial_\mu f_{\mu\nu} = B_i \bar{\eta}^i_{\nu\rho} \partial_\rho (\Delta\phi) = 0, \tag{5.88}$$

where we have made use of the antisymmetry of the 't Hooft symbols. We can therefore consider the vector field of eq. (5.84) as an abelian gauge field. Furthermore, we can attempt to double copy this field according to eq. (5.57). For the resulting graviton field to be a solution to self-dual gravity, the differential operator $\hat{k}_\mu$ must satisfy the constraints of eq. (5.21). The first constraint enforces

$$\hat{k}^2 = (A^2 + B^2)\Delta = 0 \qquad \Longrightarrow \qquad A^2 = -B^2, \tag{5.89}$$

where $B^2 = B_i B^i$, while from the second we find

$$\partial \cdot \hat{k} = A\Delta = 0. \tag{5.90}$$

Thus, for the abelian self-dual gauge field of eq. (5.84) to double copy to a valid self-dual gravity solution we fix the coefficients in the general ansatz to be

$$A = 0, \qquad B^2 = 0. \tag{5.91}$$

The components $\{B_i\}$ are therefore required to be complex. Note that these conditions are satisfied for the single copy of Eguchi-Hanson, which was given in terms of the differential operator in eq. (5.62), as would be expected. Interestingly, eq. (5.91) implies that eq. (5.79), the other constraint that arises from the 't Hooft ansatz, is automatically satisfied. That is,

$$\partial_\mu A_\mu + A_\mu A_\mu = A\Delta\phi + (A^2 + B^2)(\partial_\mu \phi)(\partial_\mu \phi) = 0. \tag{5.92}$$

Thus, the ansatz of eq. (5.84) when paired with the constraints of eq. (5.91) is of exactly the form required to construct a non-abelian SU(2) instanton via the 't Hooft ansatz, such that

$$A^a_\mu = -\bar{\eta}^a_{\mu\nu} \hat{k}_\nu \phi \tag{5.93}$$

is a solution to the full non-linear Yang-Mills equations. A potentially suggestive way to write this is as

$$A^a_\mu = -\hat{k}^a_\mu \phi, \qquad \hat{k}^a_\mu \equiv \bar{\eta}^a_{\mu\nu} \hat{k}_\nu, \tag{5.94}$$

where we have introduced a "non-abelian" differential operator $\hat{k}^a_\mu$. It is then interesting to note that

$$\hat{k}^a_\mu \hat{k}^a_\nu = \delta_{\mu\nu} \hat{k}^2 - \hat{k}_\mu \hat{k}_\nu, \tag{5.95}$$

where we have used eq. (5.77), such that

$$\hat{k}^a_\mu \hat{k}^a_\nu = -\hat{k}_\mu \hat{k}_\nu \quad \iff \quad \hat{k}^2 = 0. \tag{5.96}$$

The $\hat{k}^2 = 0$ condition is precisely that required for the fields constructed out of $\hat{k}_\mu$ to single / double copy. Furthermore, we can recognise the left-hand expression in eq. (5.96) as the operator which forms the graviton field that is the double copy of the abelian field that sits in the 't Hooft ansatz of eq. (5.93). Thus, we can in some sense double copy the non-abelian field by tracing over the colour indices of the 't Hooft symbols to construct a graviton field

$$h_{\mu\nu} = -\hat{k}^a_\mu \hat{k}^a_\nu \phi. \tag{5.97}$$

This is precisely the same gravity solution that would be obtained by double copying the abelian field of eq. (5.84). We emphasise that this is not the standard notion of the double copy. Here, instead of a replacement of colour information with kinematic information, the colour is being removed by tracing over the adjoint index. Nevertheless, an abelian and non-abelian gauge field double copying to the same gravity solution has precedence in the literature. Indeed, this was the situation discussed in chapter 3, in which abelian magnetic monopoles and their non-abelian embeddings were seen to correspond to the same gravitational solution. Also in the context of monopoles, it has previously been shown that abelian-like monopoles can be related to fully non-abelian Wu-Yang monopoles via a singular gauge transformation, thereby prompting the identification of both of these solutions as the single copy of the pure NUT solution in gravity [88]. One might then ask whether there is a gauge transformation relating the two forms of instantons considered here? This is an interesting question to which there is currently not an answer.

We have given here a general ansatz for single-copying a class of gravitational instantons, that allows us to construct both abelian and non-abelian single copies. It is then natural to ask how general this class of instanton solutions is. As discussed in section 5.1.1, instanton solutions are classified by topological invariants. In a given topological sector, a set of parameters known as collective coordinates or moduli define the space of possible instanton solutions. Some of these parameters will be redundant under e.g. gauge transformations. After factoring out such redundancies the configuration space of instanton solutions is referred to as the moduli space (see e.g. ref. [149] for an extensive review). The metric on the moduli space will not be completely smooth. It will be singular at points corresponding to instantons with zero size or where the centres of multiple instantons coincide. We can then ask: what portion of the moduli space of gauge theory instantons is covered by the ansatz of eq. (5.93)? The answer appears to be rather limited. In calculating the instanton number of eq. (5.7) for this ansatz, we find that

$$\mathrm{tr}(F_{\mu\nu}\tilde{F}_{\mu\nu}) \propto \frac{\lambda B^2}{r^8}, \tag{5.98}$$

which vanishes as is it proportional to $B^2 = 0$. The non-abelian single copy of the Eguchi-Hanson solution is therefore topologically trivial. Similar conclusions were reached for the abelian single copy in ref. [74], albeit for a different form of the field.[2] Furthermore, the single copy is a singular solution whose singularity appears non-removable via a gauge transformation, and it should therefore not be considered as part of the moduli space of gauge theory instantons. For more general scalars ϕ, similar conclusions will be reached. The instanton number will not depend on the basis chosen for the rotations generators $\{\bar{\eta}^i_{\mu\nu}\}$ entering the differential operators $\hat{k}_\mu$. It will therefore be invariant under rotations of the vector B_i, which implies that it can only depend on B^2, which is trivial. Generalisations of ϕ may still give rise to potentially interesting solutions in gravity. For example, we can consider a scalar describing N point-like disturbances at positions $\{a_i\}$:

$$\phi = \sum_{i=1}^{N} \frac{c_i}{(x-a_i)^2}. \tag{5.99}$$

Note that each a_i here is a four-vector, with i labelling the vector (i.e. not the components). Double copying a gauge theory solution generated from this scalar would appear to give a multi-centre generalisation of the Eguchi-Hanson solution. The Eguchi-Hanson solution is itself a special case of a more general class of metrics. It is equivalent to the two-centre case of the multi-centre Gibbons-Hawking metrics [161]. Thus, a graviton solution constructed from eq. (5.99) may correspond to a metric of Gibbons-Hawking type. However, we were not able to find an explicit coordinate transformation that realises this.

A further limitation of the class of instanton solutions considered here is that they are generated from a single scalar field ϕ. A more general form is that of eq. (5.25), in which the scalar ϕ^a is itself adjoint valued. It is not clear, at the level of exact classical solutions, how to turn the single scalar field ϕ into the multiple functions $\{\phi^a\}$ when taking the single copy. However, a prescription for something similar has been proposed in certain two-dimensional theories [147]. Investigating the results of this work in light of the discussion of this section will yield a number of useful insights.

[2] See ref. [22] for a discussion of the relation between these two forms of the single copy gauge field.

5.6 Insights into the kinematic algebra

5.6.1 The two-dimensional non-perturbative double copy

In ref. [147] the non-perturbative structure of the double copy in two spacetime dimensions was studied. The starting point for this work is the familiar biadjoint scalar field theory. In the conventions of ref. [147], for arbitrary gauge groups and in Lorentzian signature, the equations of motion are

$$\partial^2\phi^{aa'} - \frac{1}{2}f^{abc}\tilde{f}^{a'b'c'}\phi^{bb'}\phi^{cc'} = 0. \tag{5.100}$$

Ref. [147] proposed the following double copy rules

$$V^a \to V, \tag{5.101}$$

$$f^{abc}V^bW^c \to (\partial_\mu V)(\tilde{\partial}^\mu W), \tag{5.102}$$

where V^a and W^a are two general adjoint-valued fields. Any indices not involved in the replacements carry over unchanged. Also introduced here is the dual derivative operator

$$\tilde{\partial}^\mu = \epsilon^{\mu\nu}\partial_\nu. \tag{5.103}$$

where $\epsilon^{\mu\nu}$ is the two-dimensional Levi-Civita symbol. Applying these replacement rules to the primed indices in eq. (5.100) yields

$$\partial^2\phi^a - \frac{1}{2}f^{abc}\partial_\mu\phi^a\tilde{\partial}^\mu\phi^b = 0, \tag{5.104}$$

These are the equations of motion of Zakharov-Mikhailov (ZM) theory [162]. This is a two-dimensional theory whose equations of motion also encode the dynamics of self-dual Yang-Mills theory. Applying the double copy replacement rules a second time, now to eq. (5.104), yields

$$\partial^2\phi - \frac{1}{2}(\partial_\mu\partial_\nu\phi)(\tilde{\partial}^\mu\tilde{\partial}^\nu\phi) = 0. \tag{5.105}$$

These are the equations of motion of special Galileon (SG) theory [163–165].

The replacement rules of eqs. (5.101), (5.102) make good sense from the perspective of the double copy. The first maps an adjoint-valued field to a corresponding singlet field, while the second corresponds to the replacement of the structure constants of a colour algebra with those of a kinematic algebra. That eq. (5.102) corresponds to a replacement of structure constants can be most clearly seen in momentum space, where it is also possible to extract the explicit form of the kinematic structure constants. By Fourier transforming the right-hand side of eq. (5.102) we find

$$\begin{aligned}\int d^2x e^{ip_1\cdot x}\partial_\mu V\tilde{\partial}^\mu W &= \int d^2x \int \frac{d^2p_2}{(2\pi)^2}\int \frac{d^2p_3}{(2\pi)^2} e^{i(p_1-p_2-p_3)\cdot x}[-\epsilon^{\mu\nu}p_{2\mu}p_{3\nu}]\tilde{V}(p_2)\tilde{W}(p_3)\\ &= \int \frac{d^2p_2}{(2\pi)^2}\int \frac{d^2p_3}{(2\pi)^2}[-\epsilon^{\mu\nu}p_{2\mu}p_{3\nu}\delta^2(p_2+p_3-p_1)]\tilde{V}(p_2)\tilde{W}(p_3),\end{aligned} \tag{5.106}$$

where the momentum-space form of the fields are denoted by tildes. We can then identify the kinematic structure constants as

$$f_{p_2p_3}{}^{p_1} = X(p_2,p_3)\delta^2(p_2+p_3-p_1), \quad X(p_2,p_3) = -\epsilon^{\mu\nu}p_{2\mu}p_{3\nu}. \tag{5.107}$$

The replacement rules of eqs. (5.101), (5.102) therefore replace adjoint-valued fields with scalars, which are combined according to these kinematic structure constants. The kinematic algebra is that of area-preserving diffeomorphisms of the two-dimensional spacetime. To see this, consider the vector field

$$\boldsymbol{V} = -(\tilde{\partial}^\mu V)\partial_\mu, \tag{5.108}$$

where we use the bold-face $\boldsymbol{V}$ to distinguish the vector from the scalar V that sits within its components. An infinitesimal diffeomorphism $f^\mu\partial_\mu$ is area-preserving, or volume-preserving in higher dimensions, if $\partial_\mu f^\mu = 0$. Clearly this is the case for eq. (5.108), as eq. (5.103) implies

$$\partial_\mu\tilde{\partial}^\mu V = 0. \tag{5.109}$$

Equation (5.108) therefore corresponds to an area preserving diffeomorphism. The commutator of two diffeomorphisms yields another diffeomorphism, such that the algebra is closed:

$$[\boldsymbol{V},\boldsymbol{W}] = \boldsymbol{Z}, \qquad \boldsymbol{Z} = -\tilde{\partial}^\mu\left(\partial_\nu V\tilde{\partial}^\nu W\right)\partial_\mu. \tag{5.110}$$

To recognise the kinematic structure constants of eq. (5.107), we can take generators corresponding to individual momentum modes:

$$\boldsymbol{V}_{p_i} = -\tilde{\partial}^\mu(V_{p_i})\partial_\mu, \quad V_{p_i} = e^{ip_i\cdot x}, \tag{5.111}$$

such that

$$[\boldsymbol{V}_{p_2},\boldsymbol{V}_{p_3}] = X(p_2,p_3)\boldsymbol{V}_{p_2+p_3} = f_{p_2p_3}{}^{p_1}\boldsymbol{V}_{p_1}. \tag{5.112}$$

While it is clear that the replacement rules of eqs. (5.101, 5.102) provide a map between the three theories with equations of motion in eqs. (5.100, 5.104, 5.105), one might still ask *why* such replacements should be made. In ref. [163] the following motivation was given. It is a known fact that in the large N limit, the group $\mathrm{U}(N)$ is isomorphic to the group of diffeomorphisms of a torus [166], such that

$$\lim_{N\to\infty} \mathrm{U}(N) \sim \mathrm{Diff}_{S^1\times S^1}. \tag{5.113}$$

This can be made precise by noting that for odd N there exists a basis for the $U(N)$ generators in which they are labelled by a 2-vector whose components are integer modulo N [163, 166]. Then, in the large N limit, the $U(N)$ structure constants coincide exactly with those of the algebra of area-preserving diffeomorphisms on the torus. Returning to the three theories of eqs. (5.100, 5.104, 5.105), we note that there are two algebras present in each. Either both are the colour or kinematic algebras, or there is one of each algebra. However, as these algebras coincide for $N \to \infty$, the three theories become isomorphic in this limit. Reference [163] used this fact to argue that the colour-kinematic replacements should then also be made for finite N, and proposed a map between non-perturbative solutions to the three theories that is accurate up to sub-leading corrections in N.

5.6.2 Exact self-dual solutions and the kinematic algebra

The two-dimensional construction outlined in the previous section is highly reminiscent of the situation in self-dual Yang-Mills and self-dual gravity. Here the kinematic algebra can similarly be identified as an algebra of area-preserving diffeomorphisms, now in a certain two-dimensional plane [29]. The similarity of these two constructions can be further illuminated by rewriting the SDYM equation of motion of eq. (5.26) as

$$\partial^2\phi^a - \frac{1}{2} f^{abc}(\hat{k}_\mu\phi^b)(\partial_\mu\phi^c) = 0, \tag{5.114}$$

and the SDG equation of motion of eq. (5.24) as

$$\partial^2\phi - \frac{1}{2}(\hat{k}_\mu\hat{k}_\nu\phi)(\partial_\mu\partial_\nu\phi) = 0. \tag{5.115}$$

These expressions differ from those previously stated in that we have set the coupling constants to one and normalised the operators $\hat{k}_\mu$ such that the numerical constants are the same in both theories. Interestingly, by comparing these expressions for four-dimensional SDYM and SDG to those of two-dimensional ZM and SG theory in eqs. (5.104, 5.105) respectively, we can notice that they take precisely the same form, but with the dual derivative operator $\tilde{\partial}^\mu$ replaced with the differential operator $\hat{k}^\mu$. Indeed, we can associate $\hat{k}^\mu$ with area-preserving diffeomorphisms in an analogous manner to $\tilde{\partial}^\mu$ in two dimensions. That is, working now in four-dimensional Euclidean spacetime, we can define a general area-preserving diffeomorphism by

$$\boldsymbol{V} = -(\hat{k}^\mu \phi)\partial_\mu. \tag{5.116}$$

This is area-preserving due to the fact that, as in eq. (5.21), we have imposed

$$\partial \cdot \hat{k} = 0. \tag{5.117}$$

The cases of self-dual (SD) and anti-self-dual (ASD) Yang-Mills and gravity then arise from our general ansatz

$$\hat{k}_\mu\Big|_{\text{SD}} = \frac{1}{2} B_i \bar{\eta}^i_{\mu\nu} \partial^\nu, \quad \hat{k}_\mu\Big|_{\text{ASD}} = \frac{1}{2} B_i \eta^i_{\mu\nu} \partial^\nu. \tag{5.118}$$

The vector B_i picks out the two-dimensional plane in which the area-preserving diffeomorphisms of eq. (5.116) act. In particular, focusing on the self-dual case, we can decompose the combinations of 't Hooft symbols appearing in the differential operator of eq. (5.118) as

$$B_i \bar{\eta}^i_{\mu\nu} = b^{(1)}_{[\mu} b^{(2)}_{\nu]}, \tag{5.119}$$

where the anti-symmetrisation over indices is defined via

$$a_{[\mu} b_{\nu]} = a_\mu b_\nu - a_\nu b_\mu. \tag{5.120}$$

Using the explicit form for the 't Hooft symbols in eq. (5.76), we find the vectors $\{b^{(i)}_\mu\}$ to be

$$b^{(1)}_\mu = (B_1, B_2, B_3, 0), \quad b^{(2)}_\mu = \left(0, \frac{B_3}{B_1}, \frac{-B_2}{B_1}, -1\right), \tag{5.121}$$

where we have used that $B^2 = 0$ and assumed that $B_1 \neq 0$. The vector of eq. (5.116) can then be written as

$$\boldsymbol{V} = -\frac{1}{2}\left(b^{(1)[\mu} b^{(2)\nu]} \partial_\nu \phi\right) \partial_\mu, \tag{5.122}$$

which generates diffeomorphisms in the plane defined by the bivector $b^{(1)}_{[\mu} b^{(2)}_{\nu]}$. For example, consider the differential operator of eq. (5.62). Clearly, this corresponds to

$$B_1 = -i, \quad B_2 = 1, \quad B_3 = 0. \tag{5.123}$$

which yields

$$b^{(1)}_\mu = (-i, 1, 0, 0), \quad b^{(2)}_\mu = (0, 0, -i, -1). \tag{5.124}$$

These vectors are written here in Euclidean Cartesian coordinates. Translating to the light-cone coordinate system of eq. (5.51), we find diffeomorphisms in the (u, Y) plane

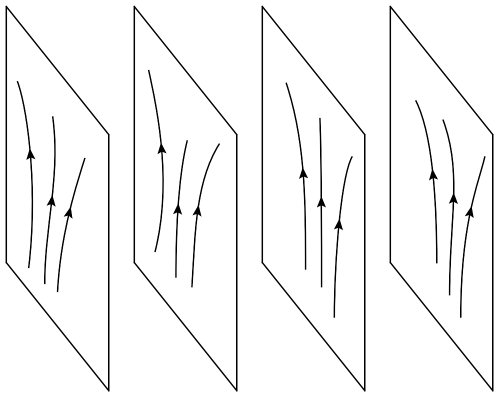


Figure 5.2: Foliation of four-dimensional Euclidean space by a family of parallel α- or β-planes, whose orientation is given by B_i. The abelian field A_μ generates area-preserving diffeomorphisms in each plane, represented here by integral curves of the vector field.

as expected. It is interesting to note that the general vectors $\{b_\mu^{(i)}\}$ in eq. (5.121) satisfy

$$b^{(i)} \cdot b^{(j)} = 0, \quad \forall i \in \{1, 2\}. \tag{5.125}$$

This makes the plane defined by the bivector $b^{(1)}_{[\mu} b^{(2)}_{\nu]}$ an example of a null plane. Null planes defined by self-dual and anti-self-dual bivectors are known as α- and β-planes respectively, and they appear frequently in the study of instanton solutions (see e.g. ref. [167]).

We have therefore found that self-dual Yang-Mills and gravity provide a four dimensional analogue of the two-dimensional non-perturbative double copy of ref. [147]. In analogy to eqs. (5.101, 5.102), we could consider the replacement rules

$$\phi^a \to \phi, \tag{5.126}$$

$$f^{abc} \phi_1^b \phi_2^c \to \partial_\mu \phi_1 \, \hat{k}^\mu \phi_2. \tag{5.127}$$

In our case, the area-preserving diffeomorphisms take place in α- or β-planes. To visualise this, we may foliate four-dimensional Euclidean space with a family of parallel α- or β-planes, whose orientation is fixed by the vector B_i. The abelian single copy field A_μ then generates area-preserving diffeomorphisms in each plane. This set-up is pictorially represented in figure 5.2. The identification of the kinematic algebra in SDYM and SDG as a certain area-preserving diffeomorphism algebra is by now a well-known fact, originally established in ref. [29]. The novel observation here is the geometric construction of these area-preserving diffeomorphisms such that they can be identified as lying in null planes with orientation controlled by the vector B_i.

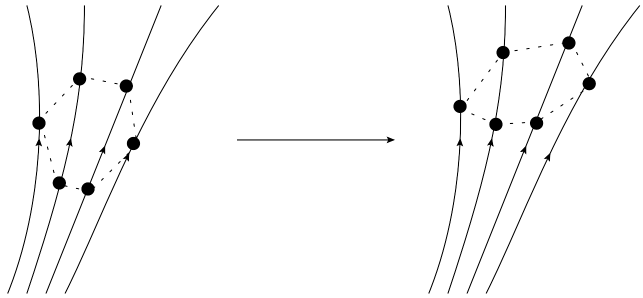


Figure 5.3: A vector field can be interpreted as defining a diffeomorphism which consists of simultaneous translations along the integral curves of the field. For the special case of an area-preserving diffeomorphism, the areas of the dotted shape on the left and right will be equal.

The discussion of this chapter is also useful for elucidating the role of the kinematic algebra at the level of exact classical solutions. For scattering amplitudes and perturbative classical solutions, the situation is clear. In the computation of an amplitude, by grouping terms based on their colour structure, one can identify factors corresponding to the two algebras present in the theory. These factors will correspond to contractions of the structure constants within each algebra. Amplitudes in each theory can then be obtained from those in another via appropriate replacements of these factors, which is permitted due to colour-kinematics duality. For perturbative classical solutions an identical situation arises by considering perturbative solutions to the momentum space equations of motion. For exact classical solutions, however, the replacement of algebras is less clear. This conceptual problem is especially pronounced given the intrinsically non-linear nature of the BCJ double copy for amplitudes. The structure constants arise in the computation of amplitudes due to the non-linear interaction terms in a given theory. In contrast, the classical solutions considered here linearise their equations of motion. With these points in mind, consider a graviton field of the form in eq. (5.57), for which the abelian-like single copy is

$$A_\mu^a = c^a A_\mu, \quad A_\mu = \hat{k}_\mu \phi. \tag{5.128}$$

We may regard a vector field A_μ as generating an infinitesimal diffeomorphism

$$A^\mu \partial_\mu. \tag{5.129}$$

This has the physical interpretation of generating a simultaneous translation along the integral curves of the field, as represented in figure 5.3. Vector fields that correspond to the single copy solutions of eq. (5.128) then generate diffeomorphisms of the form

$$(\hat{k}^\mu \phi)\partial_\mu. \tag{5.130}$$

which, due to the nature of the operator $\hat{k}_\mu$, correspond to area-preserving diffeomorphisms. Furthermore, as we have seen, these transformations can be identified as taking place in a set of null planes associated with $\hat{k}_\mu$. Thus, even for linearised solutions, there is a well-defined sense in which their properties are dictated by the kinematic algebra. In this way, we can move between the different theories related by the double copy by directly replacing the generators of one algebra with those of another. For self-dual solutions with abelian-like single copies, we find the following fields in the biadjoint, gauge, and gravity theories:

$$\Phi = (c^a T^a) \otimes (\tilde{c}^{a'} \tilde{T}^{a'})\phi, \quad A^\mu \partial_\mu = (c^a T^a)(\hat{k}^\mu \phi)\partial_\mu, \quad h^{\mu\nu}\partial_\mu\partial_\nu = (\hat{k}^\mu \hat{k}^\nu \phi)\partial_\mu\partial_\nu. \tag{5.131}$$

In moving between the fields from left to right, we can see directly the replacement of colour generators with kinematic generators of the area-preserving diffeomorphism algebra. This construction is only applicable to solutions that linearise their field equations. Interestingly, a similar restriction was found in the two-dimensional non-perturbative double copy of ref. [147], as the SG theory is ultimately related to a free theory.

5.6.3 Discussion

Equation (5.131) is simply the statement that fields in the biadjoint scalar, Yang-Mills, and gravity theories are valued in two Lie algebras. Each algebra will be either a colour or kinematic algebra, depending on which theory we are in. Unfortunately, such a construction does not straightforwardly extend to the non-abelian field of eq. (5.93), obtained via the 't Hooft ansatz. In constructing a gravitational field via eq. (5.97), it is unclear in what sense, if any, this corresponds to a replacement of algebras. We found that these gauge fields are topologically trivial and thus should not properly be considered as part of the moduli space of gauge theory instantons. The question then remains of what portion of the moduli spaces of gauge and gravity instantons are related via the double copy. In ref. [147], it was noted that all SG solutions can be used to construct solutions in ZM theory, while the opposite is not true. Thus, it may well be the case that all SDG solutions single copy to solutions in SDYM, whereas certain gauge theory instantons do not have a double copy partner.[3] A full classification of which instantons can be mapped under the double copy would be interesting. However, this

[3]Similar sentiments were expressed in ref. [168], independently of the double copy.

would almost certainly require an understanding of the exact double copy for solutions which do not linearise their equations of motion, and this remains an open problem.

Chapter 6

Perturbative insights from the self-dual sector

In the previous chapter we met self-dual Yang-Mills (SDYM) and self-dual gravity (SDG), and investigated non-perturbative solutions in these theories from the perspective of the double copy. Here we will turn our attention to the perturbative structure of these self-dual theories. SDYM and SDG are perhaps the simplest four-dimensional theories with non-trivial S-matrices. The diagrammatics of these theories mean that it is only possible to construct tree-level and one-loop amplitudes for certain fixed-helicity external states, such that SDYM and SDG are one-loop exact. Furthermore, the tree-level amplitudes are trivial, while at one-loop the amplitudes are rational functions of the external kinematics [169, 170]. Thus, the only non-trivial contribution to the S-matrices is at one-loop.

SDYM and SDG are classically integrable theories, a property that has been argued to underlie the simplicity of their S-matrices. Tree-level amplitudes are closely related to perturbative solutions to the classical equations of motion. Via an inductive construction of such solutions it is possible to show that the vanishing of the tree-level amplitudes in these theories is due to the integrability of their equations of motion [171,172]. However, it has also been suggested that this integrability is in some sense underlying the structure of the one-loop amplitudes. In ref. [171], W. Bardeen proposed that the one-loop amplitudes arise due to the anomaly of the symmetries associated with the classical integrability of the self-dual sectors.

A number of works have investigated Bardeen's idea [173–177], but it was not until recently in refs. [178–180] that a concrete realisation was developed. These works considered the uplift of the self-dual theories to twistor space. This is known to be possible for the classical theories, but is obstructed at the quantum level due to the presence of an anomaly. The authors demonstrated that via the inclusion of an additional "axion" field, the amplitudes in the self-dual theories are completely trivialised, such that the full theories are integrable.

Here we will investigate this question from the more pedestrian setting of SDYM and SDG in the light-cone gauge. The light-cone formalism is highly convenient from the perspective of scattering amplitudes; while manifest Lorentz symmetry is lost, the action is free of ghosts and contains only the propagating degrees of freedom. That is, the theory reduces to one of interacting positive and negative helicity states. This is particularly appealing in the self-dual sectors, where we are left with only a simple cubic vertex in each theory. From the perspective of the double copy, the light-cone formulation of SDYM and SDG provides a rare example of colour-kinematics duality that is manifest already in the off-shell vertices [29, 46].

We will begin by briefly reviewing the perturbative structure of SDYM and SDG in the light-cone gauge. Furthermore, to pave the way for a discussion of the quantum fate of integrability in SDYM and SDG, we will review the classical integrability of the self-dual sectors. Following this, we will introduce a quantum-corrected formalism for the self-dual theories in the light-cone gauge. By quantum-corrected we mean that quantum effects are included explicitly in the vertices via the addition of an infinite tower of one-loop effective vertices. Crucially, these effective vertices will be *loop-integrated*, in contrast to previous works that have considered effective actions for SDYM and SDG working at the level of the loop integrands [181–183]. By considering the quantum-corrected equations of motion generated by these quantum-corrected actions, we will be able to demonstrate a manifestation of Bardeen's anomaly in the spacetime formulation. To construct the explicit form of these vertices, we will take inspiration from the twistorial work of refs. [178–180] mentioned above. Initially, this approach will only work for certain restricted gauge groups due to the fact that the anomaly cancellation on twistor space occurs only for these groups. By writing the vertices in the light-cone gauge we will be able to find an $\mathrm{SU}(N)$ extension, which reproduces part of the one-loop amplitudes. We will then demonstrate that this subset of the one-loop $\mathrm{SU}(N)$ SDYM vertices double copies to the full set of one-loop SDG vertices. This double copy will follow a tree-like prescription, providing the first example of a loop-level double copy that holds at the level of the loop integrated amplitude, as opposed to the loop-integrand.

6.1 SDYM and SDG in the light-cone gauge

6.1.1 Light-cone gauge actions

We work in the light-cone coordinates $x^\mu = \{u, v, w, \bar{w}\}$, in which the Minkowski metric is

$$ds^2 = 2(-dudv + dwd\bar{w}), \tag{6.1}$$

and the wave operator takes the form

$$\Box = 2(-\partial_u\partial_v + \partial_w\partial_{\bar{w}}). \tag{6.2}$$

There are a number of standard forms of the SDYM action.[1] For our later purposes it will be most useful to start with the form [184]

$$S_{\text{SDYM}}(B, A) = \int \text{tr}(B \wedge F_{\text{ASD}}). \tag{6.3}$$

Here F_{ASD} is the anti-self-dual part of the field strength, and is valued in the Lie algebra. The B field is a Lie algebra valued anti-self-dual two form. It acts as a Lagrange multiplier in the action such that its equations of motion enforce the self-duality of F by setting

$$F_{\text{ASD}} = 0. \tag{6.4}$$

We work here in Minkowski spacetime, in which the duality conditions are

$$F = \pm i\tilde{F}, \tag{6.5}$$

with the positive and negative corresponding to a self-dual and anti-self-dual field strength respectively. In contrast to Euclidean spacetime, a gauge field satisfying the duality conditions in Minkowski spacetime is necessarily complex.

The light-cone gauge is adopted by imposing

$$A_u = 0. \tag{6.6}$$

We now follow a standard procedure to reduce the SDYM action to that of an interacting scalar theory [184]. As B is anti-self-dual, it has three independent components. Integrating out two of them sets

$$A_w = 0, \qquad \partial_u A_v = \partial_w A_{\bar{w}}. \tag{6.7}$$

[1] See refs. [184, 185] for a discussion of various actions for SDYM.

The second equation acts as an integrability condition implying

$$A_v = \frac{1}{2}\partial_w \Psi, \qquad A_{\bar{w}} = \frac{1}{2}\partial_u \Psi, \tag{6.8}$$

where Ψ is a Lie algebra valued scalar field, and the numerical coefficient is chosen for convenience. By setting the final component of B to be a second Lie algebra valued scalar $\bar{\Psi}$, we are left with the action [184]

$$S_{\text{SDYM}}(\Psi, \bar{\Psi}) = \int d^4x \, \text{tr}\left[\bar{\Psi}\big(\Box\Psi + i[\partial_u \Psi, \partial_w \Psi]\big)\right]. \tag{6.9}$$

The result is an interacting theory of two scalar fields Ψ and $\bar{\Psi}$, which can be interpreted as the positive and negative helicity degrees of freedom of the gauge field respectively. Note that variation with respect to $\bar{\Psi}$ yields the standard SDYM equations of eq. (2.46).

In SDG we have an analogous story, albeit one with more involved intermediate steps. The steps analogous to eqs. (6.6) - (6.8) yield the self-dual gravity metric [29]

$$ds^2 = 2(-du dv + dw d\bar{w}) + \partial_w^2 \phi \, dv^2 + \partial_u^2 \phi \, d\bar{w}^2 + 2\partial_u \partial_w \phi \, dv d\bar{w}, \tag{6.10}$$

and an action [186]

$$S_{\text{SDG}}(\phi, \bar{\phi}) = \int d^4x \, \bar{\phi}\big(\Box\phi + \{\partial_u \phi, \partial_w \phi\}\big), \tag{6.11}$$

where the Poisson bracket is defined as

$$\{f, g\} = \partial_u f \, \partial_w g - \partial_w f \, \partial_u g. \tag{6.12}$$

In analogy to SDYM, we may interpret ϕ and $\bar{\phi}$ as positive and negative helicity degrees of freedom respectively. Variation of this action with respect to $\bar{\phi}$ yields the Plebanski equations of self-dual gravity, given in eq. (2.47).

6.1.2 Scattering in SDYM and SDG

When talking about scattering amplitudes, we will follow a convention in which all particles in a vertex or diagram are incoming. The momentum space Feynman rules arising from the actions of eqs. (6.9, 6.11) are very simple: one only has a propagator and three-point vertex in both SDYM and SDG. The Feynman rules are:

- Propagator $(+-)$:

$$\text{SDYM:} \quad \frac{1}{k^2}\delta^{a_1 a_2} \tag{6.13}$$

$$\text{SDG:} \quad \frac{1}{k^2} \tag{6.14}$$

- Cubic vertex $(++-)$:

$$\text{SDYM:} \quad \mathcal{V}_3^{(0)}(i^+, j^+, k^-) = X(k_i, k_j) f^{a_i a_j a_k} \tag{6.15}$$

$$\text{SDG:} \quad \mathcal{V}_3^{(0)}(i^+, j^+, k^-) = X(k_i, k_j)^2 \tag{6.16}$$

- Polarisation factors for external on-shell p_i:

$$\text{SDYM:} \quad e_i^{(\pm)} = \langle \eta i \rangle^{\mp 2} \tag{6.17}$$

$$\text{SDG:} \quad e_i^{(\pm)} = \langle \eta i \rangle^{\mp 4} \tag{6.18}$$

Here we have introduced a null reference vector $\eta = |\eta\rangle[\eta|$ which defines the light-cone such that $\eta \cdot A = 0$ [46]. This makes the freedom in the choice of light-cone direction manifest and any η dependence will cancel in the computation of an amplitude. The $X(i, j)$ objects are the kinematic structure constants that we met previously in eq. (2.51). For the computation of scattering amplitudes, it will be useful to write these objects in the spinor-helicity formalism.[2] This can be done via [46]

$$X(k_i, k_j) = \langle \eta | ij | \eta \rangle. \tag{6.19}$$

As described in section 2.3.1, colour-kinematics duality and the double copy are particularly well understood in the self-dual sector. The three-point SDYM vertex of eq. (6.15) is a product of a colour structure constant $f^{a_i a_j a_k}$ and a kinematic structure constant $X(k_i, k_j)$, while the three-point SDG vertex is a product of two kinematic structure constants. Colour-kinematics duality is therefore already manifest in the off-shell vertices of SDYM and SDG.

[2] We adopt the spinor conventions of ref. [46].

With the Feynman rules in eqs. (6.13) - (6.18) it is only possible to draw the following two sets of diagrams in both SDYM and SDG:

- *Tree-level one-minus* $(- + \cdots +)$:
  n-point diagrams with one negative helicity leg and $n-1$ positive helicity legs.
- *One-loop all-plus* $(+ + \cdots +)$:
  n-point diagrams with n positive helicity legs.

No other diagrams are possible. SDYM and SDG are therefore one-loop exact, with no amplitudes at higher-loop order than one. Furthermore, the one-minus tree-level amplitudes in SDYM and SDG vanish to all n [172],

$$\mathcal{A}_n^{(0)}(1^- 2^+ \cdots n^+) = 0, \tag{6.20}$$

$$\mathcal{M}_n^{(0)}(1^- 2^+ \cdots n^+) = 0, \tag{6.21}$$

where $\mathcal{A}$ and $\mathcal{M}$ denote the amplitudes in SDYM and SDG respectively. The S-matrix in each theory is thus given only by the one-loop amplitudes. Furthermore, these one-loop amplitudes take a very simple form. They are rational functions of external kinematic data at all n. For example, the general arbitrary multiplicity form for the colour-ordered all-plus one-loop amplitudes in SDYM is [169, 187]

$$A_n^{(1)}(1^+ 2^+ \cdots n^+) = \sum_{1 \leq i_1 < i_2 < i_3 \leq n} \frac{\langle i_1 i_2 \rangle [i_2 i_3] \langle i_3 i_4 \rangle [i_4 i_1]}{\langle 12 \rangle \langle 23 \rangle \cdots \langle n1 \rangle}, \tag{6.22}$$

where we have omitted a constant multiplicative factor. A closed form expression for the one-loop amplitudes in SDG to all multiplicity also exists [170].

The possible diagrams in SDYM and SDG listed above are precisely those that can be drawn in full Yang-Mills and general relativity, such that the self-dual sectors furnish well-defined helicity sectors of the full theories. The tree-level one-minus and one-loop all-plus amplitudes in the full theories can then be computed by restricting to the self-dual sector [172, 184]. This can be made explicit by writing the actions for full Yang-Mills or general relativity in terms of interacting scalars describing positive and negative helicity degrees of freedom (see e.g. ref. [185]). The actions for the self-dual sectors then correspond to the truncation of the full actions to a single $(+ + -)$ vertex followed by a field redefinition. It is only this vertex that enters the computation of tree-level one-minus and one-loop all-plus amplitudes in the full theories.

6.1.3 Classical integrability in the self-dual sector

An important feature of the self-dual sectors of Yang-Mills and gravity is that they are classically integrable. We will here briefly review this fact for SDYM (see e.g. ref. [188] for a more extensive treatment). The story for SDG is analogous, and can be found in e.g. ref. [189].

The starting point when discussing integrability are the equations of motion. For SDYM, from the action in eq. (6.9), these are

$$0 = \Box\,\Psi + i[\partial_u\Psi, \partial_w\Psi], \tag{6.23}$$

$$0 = \Box\,\bar{\Psi} + i[\partial_u\bar{\Psi}, \partial_w\Psi] + i[\partial_u\Psi, \partial_w\bar{\Psi}]. \tag{6.24}$$

The first equation is simply the self-duality condition, arising from $F = i\tilde{F}$, and is enforced in the action via a Lagrange multiplier. A crucial feature of the second equation is that it corresponds to a linearised deformation of the first. That is, by considering a deformation of a solution to the first equation $\Psi \to \Psi + \epsilon\bar{\Psi}$ for some small ϵ, we find that $\bar{\Psi}$ solves the second equation. The solutions $\bar{\Psi}$ of the second equation therefore correspond to infinitesimal symmetries of the first.

To reveal the integrability of the classical equations of motion, consider the following two differential equations in terms of the Lie algebra valued function Λ:

$$L\Lambda = 0, \qquad M\Lambda = 0, \tag{6.25}$$

with the differential operators

$$L = \partial_u - \lambda\left(\partial_{\bar{w}} + \frac{i}{2}\partial_u\Psi\right), \qquad M = \partial_w - \lambda\left(\partial_v + \frac{i}{2}\partial_w\Psi\right), \tag{6.26}$$

where $\lambda \in \mathbb{C}P^1$ is known as the spectral parameter. The self-duality condition of eq. (6.23) then arises as the compatibility condition of this overdetermined system of equations, such that

$$[L, M] = 0 \qquad \Longleftrightarrow \qquad \Box\,\Psi + i[\partial_u\Psi, \partial_w\Psi] = 0. \tag{6.27}$$

The pair of differential operators $\{L, M\}$ in eq. (6.26) is then an example of a Lax pair. While there is no definitive method for determining whether a given system is integrable, the existence of a Lax pair is common to many integrable systems, and can thus be taken as an indication of the integrability of a theory. Consider now the expression

$$[\partial_v + \tfrac{i}{2}\partial_w \Psi, L\Lambda] - [\partial_{\bar{w}} + \tfrac{i}{2}\partial_u \Psi, M\Lambda] = \Box\,\Lambda + i[\partial_u \Lambda, \partial_w \Psi] + i[\partial_u \Psi, \partial_w \Lambda]. \tag{6.28}$$

On the right hand side we have eq. (6.24), the second equation of motion, and thus we find that Λ is a solution to this equation when it solves the differential equations in eq. (6.25). From this fact we can construct an infinite tower of solutions to the second equation of motion. Consider an expansion of Λ in terms of the spectral parameter

$$\Lambda = \sum_{r=0}^{\infty} \lambda^r \Lambda_r. \tag{6.29}$$

The overdetermined system of differential equations in eq. (6.25) is then

$$L\Lambda = \sum_{r=0}^{\infty} \lambda^r \left(\partial_u \Lambda_r - \partial_{\bar{w}} \Lambda_{r-1} - \tfrac{i}{2}[\partial_u \Psi, \Lambda_{r-1}] \right) = 0, \tag{6.30}$$

$$M\Lambda = \sum_{r=0}^{\infty} \lambda^r \left(\partial_w \Lambda_r - \partial_v \Lambda_{r-1} - \tfrac{i}{2}[\partial_w \Psi, \Lambda_{r-1}] \right) = 0, \tag{6.31}$$

with the condition that $\Lambda_{-1} = 0$. The spectral parameter is arbitrary, such that these expressions must hold for all λ. Thus, we are left with a pair of recursion relations:

$$\partial_u \Lambda_{r+1} = \partial_{\bar{w}} \Lambda_r + \frac{i}{2}[\partial_u \Psi, \Lambda_r], \qquad \partial_w \Lambda_{r+1} = \partial_v \Lambda_r + \frac{i}{2}[\partial_w \Psi, \Lambda_r]. \tag{6.32}$$

One can then recursively construct a tower of solutions $\{\Lambda_r\}$ from an initial seed solution Λ_0 to eq. (6.24). The recursion relations are compatible at level $r+1$ if Λ_r is a solution to eq. (6.24):

$$\partial_w(\partial_u \Lambda_{r+1}) - \partial_u(\partial_w \Lambda_{r+1}) = \frac{i}{2}\left(\Box\,\Lambda_r + i[\partial_u \Lambda_r, \partial_w \Psi] + i[\partial_u \Psi, \partial_w \Lambda_r]\right) = 0. \tag{6.33}$$

It then follows that Λ_{r+1} is a solution to eq. (6.24) as Ψ solves eq. (6.23):

$$\Box\,\Lambda_{r+1} + i[\partial_u \Lambda_{r+1}, \partial_w \Psi] + i[\partial_u \Psi, \partial_w \Lambda_{r+1}] = \frac{i}{2}[\Box\,\Psi + i[\partial_u \Psi, \partial_w \Psi], \Lambda_r] = 0. \tag{6.34}$$

Hence, given a solution to the second equation of motion in eq. (6.24) we can formally construct an infinite tower of additional solutions $\{\Lambda_r\}$, each of which corresponds to a linearised deformation $\Psi \to \Psi + \epsilon\Lambda_r$ of the first equation of motion in eq. (6.23).

Importantly for our purposes, this tower of infinitesimal symmetries naturally gives rise to an infinite tower of conserved currents. Note that the second equation of motion in eq. (6.24) can be expressed as the conservation of a current

$$J = \left(\partial_{\bar{w}}\bar{\Psi} + \frac{i}{2}[\partial_u\Psi, \bar{\Psi}]\right)\partial_w - \left(\partial_v\bar{\Psi} + \frac{i}{2}[\partial_w\Psi, \bar{\Psi}]\right)\partial_u, \tag{6.35}$$

such that

$$\partial_\mu J^\mu = \Box\,\bar{\Psi} + i[\partial_u\bar{\Psi}, \partial_w\Psi] + i[\partial_u\Psi, \partial_w\bar{\Psi}] = 0. \tag{6.36}$$

Let us consider this current in terms of a linearised deformation of the eq. (6.23), $\Psi \to \Psi + \epsilon\Lambda$, and expand Λ in terms of the spectral parameter as in eq. (6.29). The result is simply

$$J = \sum_{r=0}^{\infty}\lambda^r\left[\left(\partial_{\bar{w}}\Lambda_r + \frac{i}{2}[\partial_u\Psi, \Lambda_r]\right)\partial_w - \left(\partial_v\Lambda_r + \frac{i}{2}[\partial_w\Psi, \Lambda_r]\right)\partial_u\right]. \tag{6.37}$$

The current can then itself be considered as a λ-expansion, such that

$$J = \sum_{r=0}^{\infty}\lambda^r J_r, \tag{6.38}$$

where

$$J_r = \left(\partial_{\bar{w}}\Lambda_r + \frac{i}{2}[\partial_u\Psi, \Lambda_r]\right)\partial_w - \left(\partial_v\Lambda_r + \frac{i}{2}[\partial_w\Psi, \Lambda_r]\right)\partial_u. \tag{6.39}$$

With the recursion relation of eq. (6.32), we can recognise this as

$$\begin{aligned} J_r &= (\partial_u\Lambda_{r+1})\partial_w - (\partial_w\Lambda_{r+1})\partial_u \\ &= \{\Lambda_{r+1}, \cdot\}. \end{aligned} \tag{6.40}$$

The hierarchy of solutions to eq. (6.24) thus gives rise to an infinite tower of conserved currents

$$\partial_\mu J_r^\mu = 0, \quad \forall r. \tag{6.41}$$

In practice, the tower of currents is simply obtained from eq. (6.35) via the replacement

$$J_r = J(\bar{\Psi} \to \Lambda_r), \quad r = 0, 1, 2, \cdots. \tag{6.42}$$

The integrability of classical SDYM and SDG forces the tree-level amplitudes in these theories to vanish. The proof follows from the inductive construction of perturbative solutions to eq. (6.23) [171, 172]. Intuitively, the vanishing of these amplitudes follows from the fact that no amplitude can be defined that obeys the infinite tower of symmetries.

6.2 A quantum-corrected formalism

6.2.1 General quantum-corrected actions

To compute quantum effects from the SDYM and SDG actions of eqs. (6.9, 6.11) respectively, one follows the usual prescription of constructing diagrams with loops from the Feynman rules. As we have seen, SDYM and SDG are one-loop exact and so only one-loop diagrams will arise. One can then consider a quantum-corrected action by introducing an infinite set of new vertices associated with the one-loop polygon diagrams. These one-loop effective vertices will simply be off-shell diagrams with legs attached directly to the loop. With an action in this form, quantum computations proceed in a classical-like manner. That is, an n-point one-loop amplitude will be computed by considering the n-point effective vertex along with lower point effective vertices with the three-point "tree-level" vertex glued into the external legs, with all external legs taken on-shell. Furthermore, as the Feynman rules for the self-dual theories permit only all-plus diagrams at one-loop, we know that the one-loop effective vertices will have only positive-helicity external legs. We may therefore write the general structure of the quantum-corrected SDYM action as

$$S_{\text{q.c.SDYM}}(\Psi,\bar{\Psi}) = \int d^4x \, \Big(\text{tr } \bar{\Psi}\big(\Box\Psi + i[\partial_u\Psi, \partial_w\Psi]\big) + V_{\text{1-loop}}[\Psi]\Big). \tag{6.43}$$

This is simply the standard SDYM action of eq. (6.9), with a new term $V_{\text{1-loop}}$, which is suppressed by a factor of $\hbar$, corresponding to the one-loop vertices. As the one-loop vertices have only positive-helicity external legs, $V_{\text{1-loop}}$ depends only on Ψ and not $\bar{\Psi}$. It can be expanded as

$$V_{\text{1-loop}}[\Psi] = \sum_{m=2}^{\infty} V^{(m)}_{\text{1-loop}}[\Psi], \qquad V^{(m)}_{\text{1-loop}}[\Psi] \sim \Psi^m. \tag{6.44}$$

Here, $V^{(2)}_{\text{1-loop}}[\Psi]$ comes from an off-shell bubble diagram, $V^{(3)}_{\text{1-loop}}[\Psi]$ from an off-shell triangle diagram, $V^{(4)}_{\text{1-loop}}[\Psi]$ from an off-shell box diagram, and so on. By computing "tree-level" diagrams with these one-loop vertices, we will obtain the one-loop amplitudes of the theory. The situation is entirely analogous in SDG. We simply consider the standard SDG action of eq. (6.11), and add on a new term corresponding to the SDG one-loop effective vertices.

In any off-shell formulation of a theory, field redefinitions can lead to dramatically different actions, all of which give rise to the same amplitudes. The off-shell one-loop vertices in $V_{\text{1-loop}}$ are therefore non-unique. As the one-loop amplitudes in SDYM and

SDG are rational functions of external kinematics, one might expect that there are some choices of $V_{\text{1-loop}}$ in which the vertices themselves are rational (albeit non-local) functions in momentum space. We now suggest three routes to $V_{\text{1-loop}}$:

- *Momentum-space vertices from direct loop integration:*
  This is perhaps the most naive approach. Here we obtain the one-loop vertices by constructing loop level diagrams from the momentum-space light-cone gauge Feynman rules of eqs. (6.13) - (6.18) and explicitly performing the off-shell loop integrations in a given regularisation scheme. This brute-force approach is described for the first few orders in Appendix B. To summarise, we find that the two-point one-loop vertex is vanishing, $V^{(2)}_{\text{1-loop}} = 0$, with the first non-trivial vertex arising at three-points. This three-point one-loop vertex is naively quite fearsome and does not appear to be rational, as it involves dilogarithms. Upon closer inspection a rational vertex can be obtained, but in moving to higher points the computations become increasingly impracticable. This approach therefore does appear to yield rational vertices, but is not feasible in practice.

- *Single region-momenta vertex from a special regularisation of the one-loop bubble:*
  This idea has been explored in the context of SDYM in refs. [176, 177, 190]. By considering a specific regularisation for the off-shell bubble diagram, one can obtain an effective vertex from which the one-loop amplitudes arise by gluing tree-level diagrams into the external legs. This gives a simple and manifestly rational approach to the amplitude computations. However, it is crucially dependent on a region-momenta representation, which does not directly translate into a spacetime approach. We will not pursue this approach, however it would be interesting to explore how it connects to our construction.

- *Momentum-space vertices arising from anomaly cancellation in twistor space:*
  In refs. [178–180] classical SDYM and SDG were studied on twistor space. Via the introduction of a new field, it was possible to cancel an anomaly preventing the twistorial construction of these theories at the quantum-level, thereby restoring integrability to the *full*, not just the classical, theories. This renders the amplitudes in these modified forms of SDYM and SDG trivial. We shall take inspiration from these works to find a set of momentum-space vertices by integrating out the new field. More details will be given in the next section, but for now we can note two attractive properties of the vertices arising from this approach:

  (i) $V^{(m)}_{\text{1-loop}}$ are rational functions of external kinematic data.

  (ii) $V^{(m)}_{\text{1-loop}}$ are only non-trivial for $m \geq 4$, matching the fact that the one-loop amplitudes are non-trivial only for $n \geq 4$.

In the following we will focus on the third approach to the one-loop effective vertices. We will see the explicit vertices that arise from this approach in the next section. However, with the general form of the quantum-corrected action for SDYM, we can already investigate the fate of the classical integrability at the quantum level.

6.2.2 Anomalous integrability

In section 6.1.3 we reviewed the classical integrability of the self-dual theories. This was intimately tied to the classical equations of motion. At face value, this fact makes it difficult to see how one could investigate the role, if any, of this integrability in the quantum theory. However, given that our quantum-corrected action in eq. (6.43) recasts quantum computations in a classical-like manner, with this action we can attempt to proceed as we did in the classical theory.

We will again focus on SDYM, since the SDG case is analogous. From the quantum-corrected action in eq. (6.43), we can compute the quantum-corrected equations of motion:

$$0 = \Box\,\Psi + i[\partial_u \Psi, \partial_w \Psi], \tag{6.45}$$

$$0 = \Box\,\bar{\Psi} + i[\partial_u \bar{\Psi}, \partial_w \Psi] + i[\partial_u \Psi, \partial_w \bar{\Psi}] + \frac{\delta V_{\text{1-loop}}[\Psi]}{\delta \Psi}. \tag{6.46}$$

The first equation is unchanged in the quantum theory, such that it coincides with the classical equation in eq. (6.23). This is due to the fact that it is enforced by a Lagrange multiplier $\bar{\Psi}$ which does not feature in the new one-loop vertices. This first equation is still integrable; it arises as the compatibility condition of an overdetermined system of differential equations. However, in the quantum theory, the second equation of motion no longer coincides with a linearised deformation of the first, such that the classical integrability of the theory is broken at the quantum level.

Given that the vanishing of the tree-level amplitudes in the self-dual theories is a consequence of the classical integrability, it is natural to ask whether the one-loop amplitudes can be viewed as arising from the anomaly of this integrability, as first proposed in ref. [171]. In practice, we want to know what happens to the infinite tower of currents J_r associated to the classical integrability. If these currents are defined only in terms of eq. (6.45) and its linearisation, then they will continue to be conserved in the quantum theory. That is, J_r defined in terms of $\{\Lambda_r\}$, the linearised symmetries of eq. (6.45), are still conserved. However, the current J in terms of the field Ψ, given in eq. (6.35), is no longer conserved due to the presence of the one-loop vertices in eq. (6.46):

$$\partial_\mu J^\mu = \frac{1}{2}(\Box\,\bar{\Psi} + i[\partial_u\bar{\Psi}, \partial_w\Psi] + i[\partial_u\Psi, \partial_w\bar{\Psi}]) = -\frac{1}{2}\,\frac{\delta V_{\text{1-loop}}[\Psi]}{\delta\Psi}. \tag{6.47}$$

To this end, by analogy with the classical case, let us attempt to set up a quantum-corrected current, whose conservation is enforced by the second quantum-corrected equation of motion in eq. (6.46). Let us assume that the quantum-corrected action can be written in a form in which it depends on Ψ only through its derivatives, and in particular its u and w derivatives. Indeed, this is already the case for the classical action. Then, by variational integration-by-parts, we can write

$$\frac{\delta V_{\text{1-loop}}[\Psi]}{\delta\Psi} = \partial_u C^u[\Psi] + \partial_w C^w[\Psi], \tag{6.48}$$

where C^u and C^w are minus the coefficients of the variations of $\partial_u\Psi$ and $\partial_w\Psi$ respectively. While this splitting is not unique due to the presence of $\partial_u\partial_w\Psi$ terms, it is nevertheless possible. We can then write down a quantum-corrected current

$$J^{\text{q.c.}} = \left(\partial_{\bar{w}}\bar{\Psi} + \frac{i}{2}[\partial_u\Psi, \bar{\Psi}] + \frac{1}{2}\,C^w\right)\partial_w - \left(\partial_v\bar{\Psi} + \frac{i}{2}[\partial_w\Psi, \bar{\Psi}] - \frac{1}{2}\,C^u\right)\partial_u. \tag{6.49}$$

By construction, this current is conserved as a consequence of the second equation of motion of eq. (6.46), such that

$$\partial^\mu J^{\text{q.c.}}_\mu = \Box\,\bar{\Psi} + i[\partial_u\bar{\Psi}, \partial_w\Psi] + i[\partial_u\Psi, \partial_w\bar{\Psi}] + \frac{\delta V_{\text{1-loop}}[\Psi]}{\delta\Psi} = 0. \tag{6.50}$$

This is directly analogous to eq. (6.36) in the classical theory. Furthermore, in analogy with eq. (6.42), let us define

$$J^{\text{q.c.}}_r = J^{\text{q.c.}}(\bar{\Psi} \to \Lambda_r)\,, \quad r = 0, 1, 2, \cdots\,, \tag{6.51}$$

where Λ_r are the linearised symmetries of the first equation, which takes the same form in the classical and quantum theories. These quantum-corrected currents feature a universal anomaly for all r:

$$\partial^\mu J^{\text{q.c.}}_\mu = \frac{1}{2}\,\frac{\delta V_{\text{1-loop}}[\Psi]}{\delta\Psi}\,, \quad \forall r. \tag{6.52}$$

This anomaly can be thought of as generating the one-loop amplitudes. The one-loop vertices are associated with

$$\frac{\delta^m}{(\delta\Psi)^m}V_{\text{1-loop}}\Bigg|_{\Psi=0} = 2\frac{\delta^{m-1}}{(\delta\Psi)^{m-1}}\partial^\mu J^{\text{q.c.}}_\mu\Bigg|_{\Psi=0}\,, \tag{6.53}$$

and amplitudes are obtained by dressing these vertices with trees.

To summarise, we have made the breaking of the classical integrability at the quantum level manifest by defining a quantum-corrected current whose conservation is enforced by the quantum-corrected equation of motion in eq. (6.46). By considering this quantum current in terms of the linearised symmetries of the first equation of motion, we find that its conservation is blocked by the presence of the one-loop vertices. The anomaly of the classical integrability can therefore by thought of as generating the one-loop amplitudes in SDYM. We now turn to the explicit form of the one-loop vertices in SDYM and SDG.

6.3 Quantum-corrected actions for SDYM and SDG

To construct the explicit one-loop effective vertices in SDYM and SDG, we will take influence from the work of ref. [178–180]. In these works, the integrability of the self-dual theories was posited to be inherited from the ability to uplift them to local theories on twistor space. It is known that this is possible for classical SDYM and SDG, but the construction is plagued by an anomaly at the quantum level. These works then went on to modify the twistorial forms of SDYM and SDG so as to cancel the anomaly. This was done via the introduction of an "axion" which couples to $F \wedge F$ and $R \wedge R$ in SDYM and SDG respectively. Tree-level exchanges involving the axion then cancel the loop diagrams involving only gauge bosons, realising a Green-Schwartz-like anomaly cancellation [191]. This renders the S-matrices of the axion-deformed SDYM and SDG theories completely trivial, and thus integrable.

We will build off of these works by considering the spacetime theories which correspond to the axion-deformed theories on twistor space. We will use these actions to obtain a quantum-corrected action for SDYM by:

1. Integrating out the axion to obtain non-local effective vertices.
2. Flipping the sign of the effective vertices.

The sign flip in the second step follows from the fact that the amplitudes in the theories containing the axion are trivial, due to the cancellation:

$$(\textit{Loop diagrams of gauge bosons}) + (\textit{Effective diagrams for axion exchange}) = 0. \tag{6.54}$$

However, we want the effective vertices included explicitly in the action in the form of eq. (6.43). Thus, eq. (6.54) implies that we should consider

$$(\textit{Loop diagrams of gauge bosons}) = -(\textit{Effective diagrams for axion exchange}). \quad (6.55)$$

Let us now see how this works in practice. We will first consider SDYM, from which we will obtain SDG via a novel double copy.

6.3.1 Self-dual Yang-Mills

In refs. [178, 179] the following action is given

$$S_{\rho-\mathrm{SDYM}}(B, A, \rho) = \int \mathrm{tr}\Big(B \wedge F_{\mathrm{ASD}} + d^4x\, \frac{1}{2}(\Box\rho)^2 + \tilde{a}\,\rho\, F \wedge F\Big). \quad (6.56)$$

The first term is simply the action for SDYM, given in eq. (6.3). The following two terms describe the introduction of the axion ρ into the theory.[3] The second term is the kinetic term for ρ, which takes a peculiar quartic form. The third term describes the interactions, with $\tilde{a}$ a coupling constant. The cancellation between gauge field loops and tree-level axion exchanges occurs only for special restricted gauge groups. These are SU(2), SU(3), SO(8), or one of the exceptional groups [178, 179]. For each of these choices, the coupling $\tilde{a}$ is tuned to ensure the cancellation.

Let us first restrict to one of these special groups, and follow the process outlined in the two steps at the start of this section. We integrate out the axion and flip the sign of the resulting effective vertex, which yields

$$S'_{\mathrm{q.c.SDYM}}(B, A) = \int \mathrm{tr}(B \wedge F_{\mathrm{ASD}}) + d^4x\, \frac{\tilde{a}^2}{2}\Big(\frac{1}{\Box}\,\mathrm{tr}(\varepsilon^{\mu\nu\rho\lambda}F_{\mu\nu}F_{\rho\lambda})\Big)^2. \quad (6.57)$$

The prime on the action denotes that we are considering one of the restricted gauge groups. The B field appears exactly as it did in the standard action for SDYM in eq. (6.3). Thus, we can follow the same steps outlined in eqs. (6.6) - (6.8), by going to the light-cone gauge and integrating out two components of B. The result is

$$S'_{\mathrm{q.c.SDYM}}(\Psi, \bar{\Psi}) = \int d^4x\;\; \mathrm{tr}\big(\bar{\Psi}(\Box\Psi + i[\partial_u\Psi, \partial_w\Psi])\big) + a\,\Big(\frac{1}{\Box}\,\mathrm{tr}(\Psi\, \overleftrightarrow{P}^2\, \Psi)\Big)^2, \quad (6.58)$$

where the coupling a is $\tilde{a}^2$ up to a numerical factor, and is proportional to $\hbar$. We have here introduced the differential operator

$$\overleftrightarrow{P} = \overleftarrow{\partial}_u \overrightarrow{\partial}_w - \overleftarrow{\partial}_w \overrightarrow{\partial}_u, \quad (6.59)$$

[3] We note in passing that anomaly cancellation due to similar dimension-zero scalars has recently appeared in a very different context, related to the Weyl anomaly [192].

which corresponds to the kinematic structure constant of eq. (2.51) in momentum space. Note that this action is of precisely the form anticipated in eq. (6.43). The first part is the SDYM action of eq. (6.9), while the second part corresponds to $V_{\text{1-loop}}$ and depends only on Ψ and not $\bar{\Psi}$ as expected. There is, however, only a single one-loop vertex corresponding to $m = 4$. This is where the restriction of the gauge group acts. Beyond four-points there are identities among colour traces for the restricted gauge groups, which simplify the amplitudes significantly [179]. Nevertheless, it will be instructive to consider the Feynman rules derived from this action. Setting our normalisation to be $a = 1$, we have[4]

- Propagator $(+-)$:
$$\frac{1}{k^2}\delta^{a_1 a_2} \tag{6.60}$$

- $\hbar^0$-vertex $(++-)$:
$$\mathcal{V}_3^{(0)}(i^+, j^+, k^-) = X(i,j) f^{a_i a_j a_k} \tag{6.61}$$

- $\hbar^1$-vertex $(++++)$:
$$\mathcal{V}_4^{(1)}(i^+, j^+, k^+, l^+) = X(i,j)^2 \, \frac{1}{s_{ij}^2} \, X(k,l)^2 \, \delta^{a_i a_j} \delta^{a_k a_l} \tag{6.62}$$

- Polarisation factors for external on-shell p_i:
$$e_i^{(\pm)} = \langle \eta i \rangle^{\mp 2} \tag{6.63}$$

Here we have made use of the shorthand notation

$$X(k_i, k_j) = X(i,j), \qquad s_{ij\cdots n} = (k_i + k_j + \cdots k_n)^2. \tag{6.64}$$

The Feynman rules are identical to those of SDYM, but with an additional one-loop four-point vertex in eq. (6.62). A general one-loop diagram contains a single one-loop vertex and an arbitrary number of tree-level vertices. The one-loop vertex immediately yields the correct colour-ordered one-loop all-plus amplitude:

$$\Big(\prod_{i=1}^{4} \langle \eta i \rangle^{-2}\Big) \, X(1,2)^2 \, \frac{1}{s_{12}^2} \, X(3,4)^2 = \frac{[12]^2 [34]^2}{s_{12}^2} = \frac{[12][34]}{\langle 12 \rangle \langle 34 \rangle}. \tag{6.65}$$

This is the s-channel contribution. There is also a t-channel contribution, however this coincides with the former as the on-shell expression is permutation invariant in the external particles.

[4] Note that we use $\mathcal{V}$ notation for the vertex Feynman rules so as to distinguish them from the position space $V_{\text{1-loop}}$ in eq. (6.44), which contain the fields Ψ in their definition.

While this gives the colour-ordered amplitude, one might worry about the peculiar colour structure in eq. (6.62). Once again, this is where the restriction of the gauge group acts, and in ref. [179] it was explicitly confirmed that this vertex yields the full colour-dressed four-point amplitude. Proceeding to higher multiplicity, the effect of this restriction becomes more pronounced. Working with the colour-ordered vertex alone no longer yields gauge invariant quantities, let alone the desired amplitudes. This is precisely because colour structures that are independent for SU(N) are no longer independent for the special restricted gauge groups. Thus, while the colour-dressed amplitudes will of course be gauge invariant, the naive colour-ordered amplitudes will not be.

To extend the above Feynman rules to SU(N) we will need to introduce additional vertices. To this end, it will be useful to revisit the general form for the colour-ordered one-loop all-plus amplitudes in SDYM, written in eq. (6.22). When first constructed in ref. [169], this was observed to split such that

$$A_n^{(1)}(1^+2^+\cdots n^+) = \frac{E(123\cdots n) + O(123\cdots n)}{\langle 12\rangle\langle 23\rangle\cdots\langle n1\rangle}, \tag{6.66}$$

where we have omitted an overall normalisation constant and the E- and O-parts are given by

$$E(123\cdots n) = \sum_{1\le i_1<i_2<i_3<i_4\le n} \langle i_1 i_2\rangle[i_2 i_3]\langle i_3 i_4\rangle[i_4 i_1] + [i_1 i_2]\langle i_2 i_3\rangle[i_3 i_4]\langle i_4 i_1\rangle \tag{6.67}$$

$$O(123\cdots n) = \sum_{1\le i_1<i_2<i_3<i_4\le n} \langle i_1 i_2\rangle[i_2 i_3]\langle i_3 i_4\rangle[i_4 i_1] - [i_1 i_2]\langle i_2 i_3\rangle[i_3 i_4]\langle i_4 i_1\rangle \tag{6.68}$$

In the O-part we can recognise the presence of the four-dimensional Levi-Civita symbol

$$\begin{aligned}\varepsilon(i,j,k,l) &= 4i\varepsilon_{\mu\nu\rho\sigma}k_i^\mu k_j^\nu k_k^\rho k_l^\sigma \\ &= [ij]\langle jk\rangle[kl]\langle li\rangle - \langle ij\rangle[jk]\langle kl\rangle[li],\end{aligned} \tag{6.69}$$

such that

$$O(123\cdots n) = -\sum_{1\le i_1<i_2<i_3<i_4\le n} \varepsilon(i_1,i_2,i_3,i_4). \tag{6.70}$$

There are two classes of vertices needed to extend the one-loop vertices to SU(N): those that extend the four-point vertex of eq. (6.62) and those containing the Levi-Civita symbol. In the on-shell limit, the first class will yield the E-part of the amplitude, while the second class will give the O-part. Notice that the O-part vanishes at four-points, and thus was not required in eq. (6.65).

In the following we will consider the colour-dressed amplitudes

$$\mathcal{A}_n^{(1)} = \sum_{\sigma \in S_{n-1}} c^{a_{\sigma(1)}a_{\sigma(2)}a_{\sigma(3)}\cdots a_{\sigma(n)}}\, A_n^{(1)}\big(\sigma(1)\sigma(2)\sigma(3)\cdots\sigma(n)\big), \tag{6.71}$$

where the sum is over all non-cyclic permutations. Here we have introduced the notation

$$c^{a_1a_2a_3\cdots a_n} = f^{b_1a_1b_2} f^{b_2a_2b_3} f^{b_3a_3b_4} \cdots f^{b_na_nb_1} \tag{6.72}$$

for the cyclic n-gon colour factors. It will also be useful to introduce symmetrisation and antisymmetrisation over the indices respectively as

$$c^{(a_1a_2)a_3\cdots a_n} = c^{a_1a_2a_3\cdots a_n} + c^{a_2a_1a_3\cdots a_n} \tag{6.73}$$

$$c^{[a_1a_2]a_3\cdots a_n} = c^{a_1a_2a_3\cdots a_n} - c^{a_2a_1a_3\cdots a_n}. \tag{6.74}$$

When nested (anti)symmetrisations appear, they are evaluated sequentially from the outer-most moving inwards, such that e.g.

$$c^{[(a_1a_2)a_3]a_4\cdots a_n} = c^{(a_1a_2)a_3a_4\cdots a_n} - c^{a_3(a_1a_2)a_4\cdots a_n}. \tag{6.75}$$

By trial-and-error, we were able to find the following set of m-point one-loop SU(N) vertices, which reproduce the E-part of the amplitude:

- $\hbar^1$-vertices $(++\cdots+)$:

$$\mathcal{V}_m^{(1)}(1^+,2^+,\ldots,m^+) = \sum_{i=2}^{m-2} X(k_1,k_2)^2 \left(\prod_{j=2}^{i-1} \frac{X(k_{1,\ldots,j}\,,k_{j+1})}{s_{1\cdots j}}\right) \frac{1}{s_{1\cdots i}^2} \left(\prod_{l=i+1}^{m-2} \frac{X(k_{1,\ldots,l-1}\,,k_l)}{s_{1\cdots l}}\right) X(k_{m-1},k_m)^2 \cdot c^{[[\cdots[(a_1a_2)a_3]\cdots]a_i][a_{i+1}[\cdots[a_{m-2}(a_{m-1}a_m)]\cdots]]}. \tag{6.76}$$

Here we have defined the notation

$$k_{i,j,\ldots,n} = k_i + k_j + \cdots + k_n. \tag{6.77}$$

Importantly, the symmetry/antisymmetry of the colour factors under interchange of indices is mimicked by the kinematic numerators. The full set of Feynman rules to compute the E-part of the amplitude to n-points is then the rules in eqs. (6.60 - 6.63), but with the four-point one-loop vertex in eq. (6.62) replaced by this set of general m-point one-loop vertices. We have checked that these vertices give the E-part of the one-loop amplitudes numerically up to seven-points.

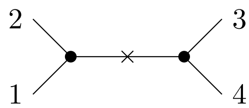
At first glance, eq. (6.76) is not a particularly appealing expression. To make the structure of these vertices clear, let us write the first few orders in m explicitly:

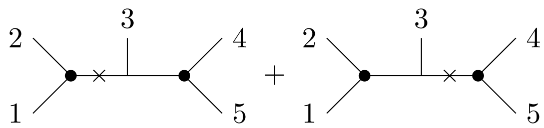
$$\mathcal{V}_4^{(1)} = \frac{X(1,2)^2 X(3,4)^2}{s_{12}^2} c^{(a_1a_2)(a_3a_4)}, \tag{6.78}$$

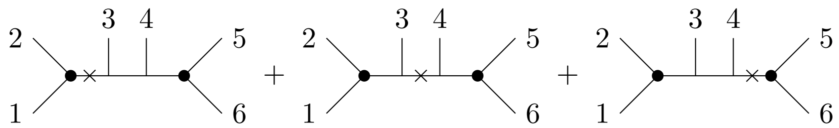
$$\begin{aligned} \mathcal{V}_5^{(1)} = &\frac{X(1,2)^2 X(3,4+5) X(4,5)^2}{s_{12}^2 s_{45}} c^{(a_1a_2)[a_3(a_4a_5)]} \\ &+ \frac{X(1,2)^2 X(1+2,3) X(4,5)^2}{s_{12} s_{45}^2} c^{[(a_1a_2)a_3](a_4a_5)}, \end{aligned} \tag{6.79}$$

$$\begin{aligned} \mathcal{V}_6^{(1)} = &\frac{X(1,2)^2 X(3,4+5+6) X(4,5+6) X(5,6)^2}{s_{12}^2 s_{123} s_{56}} c^{(a_1a_2)[a_3[a_4(a_5a_6)]]} \\ &+ \frac{X(1,2)^2 X(1+2,3) X(4,5+6) X(5,6)^2}{s_{12} s_{123}^2 s_{56}} c^{[(a_1a_2)a_3][a_4(a_5a_6)]} \\ &+ \frac{X(1,2)^2 X(1+2,3) X(1+2+3,4) X(5,6)^2}{s_{12} s_{123} s_{56}^2} c^{[[(a_1a_2)a_3]a_4](a_5a_6)}. \end{aligned} \tag{6.80}$$

Here we have used momentum conservation in the kinematic factors to make the shared symmetries with the colour factors clear. To further simplify matters, we can introduce some diagrammatic notation for these vertices:

$\mathcal{V}_4^{(1)} =$  , (6.81)

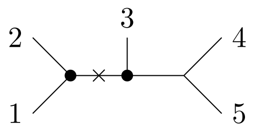
$\mathcal{V}_5^{(1)} =$  , (6.82)

$\mathcal{V}_6^{(1)} =$  . (6.83)

Here the one-loop vertices are given a representation as a dressed line. A standard vertex indicates an X, while a bold vertex indicates an X^2. Furthermore, a standard propagator indicates an s, while a cross on the propagator indicates an s^2. The colour factors of each diagram are then simply read off so as to reflect the symmetry/antisymmetry of the factors of X and X^2. From these diagrams the general structure of the one-loop effective vertices in eq. (6.76) is clear. For each increasing m, we attach a new external leg to the internal line, and consider a contribution from every possible position of the squared propagator. It is important to emphasise that while these diagrams look like trees they are truly representing one-loop vertices.

With this diagrammatic notation in hand it is easy to represent the structure of the one-loop vertices at any m. Furthermore, it extends to diagrams formed from the combination of a one-loop vertex and the tree-level vertex of eq. (6.61). For all $m > 4$, the E-part of the amplitude will receive contributions from the m-point vertex as well as all lower-point one-loop vertices dressed with trees. For example, a contribution of the four-point vertex at five-points is

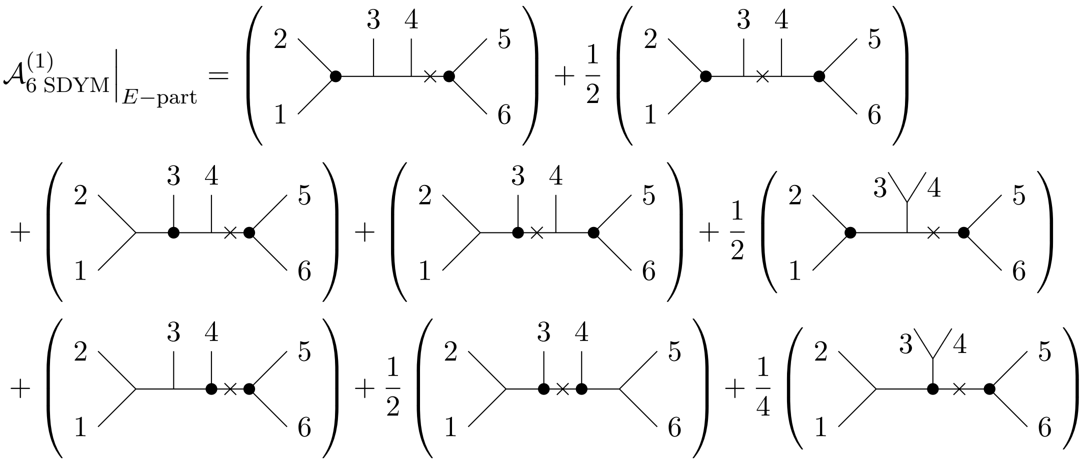
$$\begin{aligned}
&\mathcal{V}_4^{(1)}(1^+, 2^+, 3^+, (4+5)^+) \cdot \mathcal{V}_3^{(0)}(4^+, 5^+, -(4+5)^-) \\
&\qquad = X(1,2)^2 \, \frac{1}{s_{12}^2} \, X(3, 4+5)^2 \, c^{(a_1 a_2)(a_3 b)} \cdot \frac{X(4,5)}{s_{45}} \, f^{b a_4 a_5} \\
&\qquad = \frac{X(1,2)^2 X(3,4+5)^2 X(4,5)}{s_{12}^2 s_{45}} c^{(a_1 a_2)(a_3 [a_4 a_5])} \\
&\qquad =
\end{aligned}$$



. (6.84)

This contribution is formed from gluing the three-point tree-level vertex into the four-point one-loop vertex, where we have used a dot in the first and second lines to clearly demarcate the contribution from each vertex. After combining the colour factors, the result can be succinctly written diagrammatically, following precisely the same rules as for the one-loop vertices themselves. A more involved example is the full E-part of the six-point amplitude, which is given by

$$\mathcal{A}^{(1)}_{6\ \mathrm{SDYM}}\Big|_{E-\mathrm{part}} =$$



$$+ \ \text{permutations}\{123456\} \,. \tag{6.85}$$

Here the first, second, and third lines contain the contributions from the one-loop vertices of multiplicity six, five, and four respectively. The numerical coefficients are symmetry factors compensating for overcounting, which allows us to write the full result in a more compact form.

In eq. (6.76) we have given the Feynman rules for a set of m-point one-loop effective vertices. The natural question then is how these vertices are encoded within the action of eq. (6.43). This is done via the inclusion of

$$V_{\text{1-loop}}[\Psi]\Big|_{E-\text{part}} = \left((\Psi \overset{\leftrightarrow}{P}{}^2 \Psi) \frac{1}{\overset{\leftarrow}{\Box} - \overset{\leftarrow}{\text{ad}}_{\Psi \overset{\leftrightarrow}{P}}} \right)^{bc} \left(\frac{1}{\overset{\rightarrow}{\Box} - \overset{\rightarrow}{\text{ad}}_{\Psi \overset{\leftrightarrow}{P}}} (\Psi \overset{\leftrightarrow}{P}{}^2 \Psi) \right)^{cb}, \tag{6.86}$$

where $\text{ad}_X Y = [X, Y]$ and we have the geometric series

$$\begin{aligned} \left(\frac{1}{\overset{\rightarrow}{\Box} - \overset{\rightarrow}{\text{ad}}_{\Psi \overset{\leftrightarrow}{P}}} (\Psi^{a_1} \overset{\leftrightarrow}{P}{}^2 \Psi^{a_2}) T^{a_1} T^{a_2} \right)^{bc} &= \frac{1}{\overset{\rightarrow}{\Box}} f^{b(a_1|e} f^{e|a_2)c} (\Psi^{a_1} \overset{\leftrightarrow}{P}{}^2 \Psi^{a_2}) \\ &+ (f^{ba_3d} f^{d(a_1|e} f^{e|a_2)c} - f^{b(a_1|e} f^{e|a_2)d} f^{da_3c}) \frac{1}{\overset{\rightarrow}{\Box}} (\Psi^{a_3} \overset{\leftrightarrow}{P} \frac{1}{\overset{\rightarrow}{\Box}} (\Psi^{a_1} \overset{\leftrightarrow}{P}{}^2 \Psi^{a_2})) \\ &+ \mathcal{O}\Big(ffff \frac{1}{\Box} (\Psi P \frac{1}{\Box} (\Psi P \frac{1}{\Box} (\Psi P^2 \Psi))) \Big). \end{aligned} \tag{6.87}$$

In the final term we have indicated schematically the next term in the series. Let's examine the second line. The colour structure will contribute towards a factor of $c^{\cdots[a_3(a_1 a_2)]}$, while the $\overset{\leftrightarrow}{P}$ operators will contribute towards a kinematic numerator $[\cdots X(k_3, k_1 + k_2) X(k_1, k_2)^2]$. This is as expected from eq. (6.76). Furthermore, the squared Mandelstam variable arises from the $\Box^{-2}$ that is generated from the $\overset{\leftarrow}{\Box}{}^{-1}$ and $\overset{\rightarrow}{\Box}{}^{-1}$ in the left and right brackets of eq. (6.86).

Our one-loop effective vertices reproduce the E-part of the amplitude. We were unable to obtain the O-part. No such vertex exists at four-points, while at five-points a valid vertex can be obtained by taking the O-part of the amplitude off-shell:

$$\frac{X(k_1, k_2)}{s_{12}} \frac{X(k_2, k_3)}{s_{23}} \frac{X(k_3, k_4)}{s_{34}} \frac{X(k_4, k_5)}{s_{45}} \frac{X(k_5, k_1)}{s_{51}} \varepsilon(1, 2, 3, 4)\, c^{a_1 a_2 a_3 a_4 a_5}. \tag{6.88}$$

Here the cyclic product of $\langle ij \rangle^{-1}$ factors is continued off-shell into a cyclic product of $X(i, j)/s_{ij}$ factors. When the polarisation factors are included, the $|\eta\rangle$ dependence introduced by X is eliminated. We were, however, unable to find a general form for these vertices. Thus, a closed form for the complete quantum-corrected SU(N) SDYM action still remains an open problem. Despite this, it turns out that it is only the E-part that is required to obtain SDG via the double copy.

6.3.2 Self-dual gravity

In ref. [180] the gravitational counterpart of the twistorial construction of refs. [178, 179] was developed. Here we will follow a different route, and obtain the quantum-corrected action for SDG via the double copy. Recall, that the self-dual sector provides a rare off-shell example of colour-kinematics duality, as first demonstrated in ref. [29]. Furthermore, this has been shown to extend to loop-level in the self-dual theories, applying to the loop-integrands of the one-loop amplitudes [46]. This is the standard story for loop-level examples of the double copy; the duality acts at the level of the loop-integrand. There is at present no indication that colour-kinematics duality extends to the final expression for the amplitude, after loop integration. This is the surprise we find here. We will see that a tree-like double copy prescription applies to the loop-integrated one-loop effective vertices, thereby allowing us to simply obtain the corresponding vertices in SDG.

The quantum-corrected action for SDG takes a directly analogous form to that in SDYM, given in eq. (6.43). We take the SDG action in light-cone gauge and add on an infinite set of one-loop effective vertices:

$$S_{\text{q.c.SDG}}(\phi,\bar{\phi}) = \int d^4x \,\Big(\bar{\phi}\big(\Box\phi + \{\partial_u\phi, \partial_w\phi\}\big) + V_{\text{1-loop}}[\phi]\Big). \tag{6.89}$$

The Feynman rules arising from this action will be those of SDG, listed in eqs. (6.13) - (6.18), along with a set of $\hbar^1$ vertices from $V_{\text{1-loop}}[\phi]$. To obtain the one-loop vertices, we follow a tree-like double copy prescription. We take the one-loop SDYM vertices of eq. (6.76), replace the colour numerators with a second copy of the kinematic numerators, and leave the denominators untouched. The result is:

- $\hbar^1$-vertices $(++\cdots+)$:

$$\mathcal{V}^{(1)}_m(1^+,2^+,\dots,m^+) =$$
$$\sum_{i=2}^{m-2} X(k_1,k_2)^4 \left(\prod_{j=2}^{i-1} \frac{X(k_{1,\dots,j}\,,k_{j+1})^2}{s_{1\cdots j}}\right)\frac{1}{s^2_{1\cdots i}}\left(\prod_{l=i+1}^{m-2}\frac{X(k_{1,\dots,l-1}\,,k_l)^2}{s_{1\cdots l}}\right)X(k_{m-1},k_m)^4. \tag{6.90}$$

We have verified numerically that the amplitudes computed from the effective vertices written here reproduce the one-loop all-plus SDG amplitudes up to six-points. The n-point expression for these amplitudes can be found in ref. [170]. Furthermore, these vertices can be simply encoded within the action of eq. (6.89) via the inclusion of

$$V_{\text{1-loop}}[\phi] = \Big(\frac{1}{\Box_\phi} \, (\phi \overset{\leftrightarrow}{P}{}^4 \phi) \Big)^2 = \left(\frac{1}{\Box - (\phi \overset{\leftrightarrow}{P}{}^2 \,\cdot)} \, (\phi \overset{\leftrightarrow}{P}{}^4 \phi) \right)^2 , \tag{6.91}$$

where $\Box_\phi$ is the wave operator evaluated on the background of eq. (6.10). Perhaps surprisingly, this term is far simpler than its SDYM counterpart in eq. (6.86), which itself only gave rise to the E-part of the amplitude.

As a consistency check of the discussion in this section, we note that the quantum-corrected SDG action of eqs. (6.89, 6.91) can be written in the covariant form

$$S_{\text{q.c.SDG}} = S_{\text{SDG}} + b \int d^4x \sqrt{|g|} \left(\frac{1}{\Box_g} \left(\frac{\varepsilon^{\rho\lambda\sigma\omega}}{\sqrt{|g|}} \, R^\mu{}_{\nu\rho\lambda} \, R^\nu{}_{\mu\sigma\omega} \right) \right)^2 . \tag{6.92}$$

Here b is a normalisation constant of the one-loop amplitudes, and $\Box_g$ is the wave operator on a curved background with metric $g_{\mu\nu}$. This action arises from the anomaly-free theory introduced in ref. [180],

$$S_{\rho-\text{SDG}} = S_{\text{SDG}} + \int \left(d^4x \sqrt{|g|} \, \frac{1}{2} (\Box_g \, \rho)^2 + \tilde{b} \, \rho \, R^\mu{}_\nu \wedge R^\nu{}_\mu \right) , \tag{6.93}$$

where $b = \frac{1}{2}\tilde{b}^2$, after one integrates out the axion ρ and flips the sign of the resulting interaction term, as prescribed at the start of this section. Therefore, the result of ref. [180] already encodes the quantum-corrected action of SDG. We have arrived at the same answer by extending the SDYM vertices to SU(N) and performing a double copy.

6.3.3 Discussion

In this section we have found a set of fully-integrated one-loop effective vertices in SDYM which double copy to give effective vertices in SDG. The SDYM vertices are not the full set of vertices in the theory, but rather reproduce the E-part of the one-loop SDYM amplitudes in eq. (6.66). This misses the O-part, and an understanding of this class of vertices is needed to define the full quantum-corrected action of SDYM. Furthermore, it is unclear why only the E-part of the amplitude is necessary to obtain the SDG vertices via the double copy. Interestingly, in the paper where the n-point one-loop SDG amplitudes were first constructed, ref. [170], it was noticed that when written in a particular way there is a similarity between the E-part of the SDYM amplitude and the full SDG amplitude. Here we have shown that this relationship takes the form of a unique loop-level double copy, that applies at the level of the loop-integrated amplitudes. Investigating the absence of the O-part of the amplitude in this correspondence may shed light on a broader understanding of double copies of this form.

At a more practical level, it would be interesting to examine how the anomaly affects the full theories. One could imagine replacing the first term in the action of eq. (6.57) with the full Yang-Mills action. Going to the light-cone gauge and integrating out auxiliary components of the gauge field would then give the Yang-Mills action in the Chalmers-Siegel form [185] with a set of additional one-loop vertices. These effective vertices would now feature additional helicity configurations, as opposed to just the all-plus vertex found in SDYM. These would certainly give rational contributions and thus it would be interesting to check whether they give *the* rational parts of their respective amplitudes. Turning to gravity, it has been observed that the two-loop divergence in pure general relativity is connected to the one-loop like-helicity amplitudes [193]. One could then ask what happens if we explicitly cancel the anomaly via the inclusion of our effective vertices in the Einstein-Hilbert action? Testing whether this eliminates the divergence would be a daunting task, but an interesting one nonetheless.

Chapter 7

Conclusions

Colour-kinematics duality and the double copy provide a surprising link between different field theories. When viewed through these lenses, gauge theory and gravity appear to be far more closely related than traditional approaches suggest. This is a startling fact given their physical differences, but offers a tantalising glimpse of a more general framework within which to view these theories. In this work we have studied a number of novel manifestations of the double copy, both extending the known catalogue of examples in which it is known to apply and providing insights into the underlying structure of the correspondence.

A global view of the double copy

In chapter 3 we extended the double copy between magnetic monopole solutions in gauge theory and the Taub-NUT solution in general relativity. These solutions have previously been related via a *local* double copy, in which the gauge field at a given spacetime point is identified as the single copy of the graviton field at the same point [70, 88]. Here we generalised this local structure to arbitrary gauge groups by showing that non-abelian monopole gauge fields can always be written in a form that consists of a dressed abelian-like Dirac monopole. The double copy of both abelian and non-abelian monopole fields then corresponds to the pure NUT solution in gravity, as was found for SU(2) gauge theory in ref. [88].

Next we promoted this local story to a global one, by identifying a topological characterisation that can be matched up on either side of the double copy in analogy to the local structure of these solutions. Finding such a correspondence at first seems problematic, as the non-trivial topology of solutions is typically characterised by topological invariants, which crucially depend on the gauge group. As the double copy for magnetic monopoles seems indifferent to the specifics of the gauge group, it is then unclear what the relevant topological quantity is. Here we showed that the relevant topological quantity is a patching condition between gauge or graviton fields defined in different domains. The introduction of more than one field in each theory is necessary to define a non-singular field configuration everywhere away from the origin. However, to do this in a consistent manner, one necessarily induces a non-trivial topology which can be neatly encapsulated in the patching conditions of eqs. (3.79, 3.98). The form of the patching condition in gauge theory is independent of the gauge group, and it takes a precisely analogous form in gravity. Monopoles and the Taub-NUT metric therefore constitute an example of exact solutions for which local and global structure can be related via the double copy.

Wilson lines and holonomy

In chapter 4 we developed an understanding of the double copy structure of Wilson lines and holonomy. We began by identifying the double copy of the gauge theory Wilson line to be the gravitational operator of eq. (4.41). This operator has been labelled as a gravitational Wilson line before in e.g. refs. [112,114–117] due to the analogous physical role that it plays to the gauge theory Wilson line. Here we show that this analogy takes the form of a double copy. This proposal, however, raised a difficult question. When the gauge theory Wilson line is integrated over a closed loop, it corresponds to the holonomy, however its double copy is *not* the gravitational holonomy operator. What then is the single copy of the gravitational holonomy?

To answer this question we discussed how the gravitational holonomy arises naturally in the point particle action for a spinning compact object, where it describes the interaction of a spinning test particle with a non-trivial gravitational background. As the single copy of such actions has been previously identified from a perturbative analysis of the radiation emitted from the spinning source, we were able to identify the single copy of the gravitational holonomy to be eq. (4.80). This operator, which we dubbed the SCH operator, arises as a spin-dependent correction to the standard gauge theory Wilson line. To further justify this double copy, we discussed the role the operator plays in the IR structure of gauge and gravity amplitudes, and studied the form it takes for the algebraically special Kerr-Schild solutions.

The SCH operator forms the SCH group. To conclude this chapter we studied the form of the holonomy and SCH groups for a collection of well-known solutions that are double copies of one another. Here we found three distinct cases. For the Schwarzschild solution, both groups retain their maximal form, while for self-dual solutions they reduce to mutually isomorphic subgroups. For the case of magnetic monopoles and the pure NUT solution, the SCH group was found to reduce while the holonomy group did not. While this mismatch appears unexpected, we explained how it can be traced back to the inherent physical differences between these two solutions.

Non-perturbative aspects of the self-dual double copy

In chapter 5 we investigated the non-perturbative structure of the double copy by restricting to the self-dual sectors of pure Yang-Mills and general relativity. We began by searching for exact power-like spherically symmetric solutions to the equations of motion in Euclidean biadjoint scalar theory, in the hope that such solutions may relate to known instantons in gauge theory and gravity. We found that while a general d-dimensional solution exists, it vanishes in four-dimensions. This was due to the fact that the power-like form that we found is a harmonic function in $d = 4$, such that it solves the linearised biadjoint field equations. Furthermore, this linearised solution can be identified as the zeroth copy of the Eguchi-Hanson solution, a well known gravitational instanton in general relativity.

An exact Kerr-Schild-like double copy exists for self dual solutions, as reviewed in section 5.1.3. Here the graviton is written in terms of a pair of differential operators, such that it decomposes into a local product in momentum space. We interpreted the differential operators which appear in the Eguchi-Hanson solution as a specific example of a more general ansatz, involving the 't Hooft symbols which are common in the study of instantons. Abelian gauge fields constructed in terms of our ansatz are of precisely the form that is required to form non-abelian instanton solutions via the 't Hooft ansatz. This fact meant that we could immediately identify a truly non-abelian instanton solution with a given abelian-like single copy of a self-dual gravity solution. This is highly reminiscent of the situation for magnetic monopoles [1, 88].

Our general ansatz for the differential operators appearing in the exact self-dual double copy also provided useful insights into the nature of the kinematic algebra. It allowed for a four-dimensional generalisation of the two-dimensional non-perturbative double copy proposed in ref. [147]. Furthermore, we discussed a geometric interpretation of the ansatz, in which it controls the null planes in which the kinematic area-preserving diffeomorphism algebra acts. From this observation we were able to demonstrate how to interpret the replacement of colour with kinematic algebras at the level of exact classical solutions.

An old anomaly and a new double copy in quantum SDYM and SDG

Finally, in chapter 6 we investigated an old idea from W. Bardeen [171], that the one-loop amplitudes in self-dual Yang-Mills and self-dual gravity are generated by the anomaly of the classical integrability of these theories. To this end, we developed explicit quantum-corrected actions for SDYM and SDG. These actions take the form of the standard light-cone gauge action for each theory with an infinite number of one-loop effective vertices added on. With this general form we were able to demonstrate a manifestation of Bardeen's idea, by defining a quantum-corrected current whose conservation corresponded to the equation of motion that is modified in the quantum theory. Considering this quantum current in terms of the symmetries associated with the classical integrability, we saw that the conservation of the current was blocked by the presence of the one-loop effective vertices.

To construct the explicit loop-integrated form for the one-loop vertices we took inspiration from recent work in twistor space [178–180]. Here modified actions for SDYM and SDG were presented whose amplitudes are trivial due to a cancellation between tree-level "axion" diagrams and loop-level gauge boson diagrams. To obtain our vertices we integrated out the axion field and flipped the sign of the resulting effective vertex. The one-loop amplitudes could then be computed by proceeding in a tree-like manner, with diagrams formed from the gluing together of one-loop and tree-level vertices. We then extended the vertices obtained in this way to full SU(N) SDYM, and showed that this set of vertices double copies to SDG. This was an unexpected result. In all previous incarnations of the double copy at loop-level, it holds at the level of the loop-integrand. Here we find that it holds at the level of the loop-integrated effective vertices of the quantum-corrected actions.

Concluding remarks

In this thesis we have probed the double copy from a variety of different perspectives. We hope that these results provide a step, however small, along the road to a full understanding of this fascinating correspondence. Many questions remain unanswered. Can the correspondence be promoted to a general statement at loop-level? What is the true nature of the kinematic algebra? How generally can we manifest this structure at the classical level? It is unclear when, how, or even if answers to these questions will appear. However, as the archive of known double copies continues to expand and new insights into the theoretical structure of the correspondence are unearthed, one can hope that glimpses of a more general framework might gradually emerge.

Appendix A

Derivation of the Kerr-Schild spin connection

In this appendix, we provide a derivation of the spin connection in Kerr-Schild coordinates. The spin connection satisfies Cartan's first structure equation in the absence of torsion,

$$de^a + \omega^a{}_c \wedge e^c = 0. \tag{A.1}$$

In tensorial language this takes the form

$$\partial_\mu e^a{}_\nu - \partial_\nu e^a{}_\mu + (\omega_\mu)^a{}_\nu - (\omega_\nu)^a{}_\mu = 0, \tag{A.2}$$

where we have contracted the vielbein with the spin connection. Multiplying through by a factor of $e_b{}^\mu e_a{}^\nu$ then yields

$$(\partial_b e^a{}_\nu) e_a{}^\nu - (\partial_b e^a{}_\mu) e_b{}^\mu + (\omega_b)^a{}_c - (\omega_c)^a{}_b = 0. \tag{A.3}$$

Next, we can substitute the explicit forms of the Kerr-Schild vielbein given in eq. (4.92), and use the null condition from eq. (2.62), to obtain

$$\partial_b e^a{}_c - \partial_c e^a{}_b + (\omega_b)^a{}_c - (\omega_c)^a{}_b - \frac{1}{4}\phi^2 k^a k^\mu \left[k_c \partial_b k_\mu - k_b \partial_c k_\mu\right] = 0. \tag{A.4}$$

Upon lowering the index a, we can cyclically permute the indices (a, b, c) and consider the combination $(a, b, c) - (b, c, a) - (c, a, b) = 0$, to find

$$(\omega_\mu)_{bc} = (\omega_a)_{bc} e^a{}_\mu = (\partial_c e_{ab} - \partial_b e_{ac})\, e^a{}_\mu + \frac{1}{4}\phi^2 k_\mu k^\nu \left(k_c \partial_b k_\nu - k_b \partial_c k_\nu\right), \tag{A.5}$$

where we have also multiplied the entire equation by ${e^a}_\mu$. Finally, we again make use of eqs. (4.92) and (2.62) for the vielbein in the first term, after which cancellations occur and we are left with

$$(\omega_\mu)_{ab} = \partial_b e_{a\mu} - \partial_a e_{b\mu}. \tag{A.6}$$

This expression agrees with the similar result found in ref. [194]. Note also that an alternative route to this expression is to substitute the explicit forms for the Kerr-Schild vielbein of eq. (4.92) into eq. (4.22).

Appendix B

One-loop vertices from off-shell loop integration

In this appendix we briefly illustrate the brute-force approach to computing the one-loop effective m-point vertices in SDYM and SDG. For each m, this involves computing the one-loop diagram with m off-shell external legs directly attached to the loop by using the light-cone gauge Feynman rules in eqs. (6.13) - (6.18) and explicitly performing the loop integration in a given regularisation scheme. Here, we discuss this procedure for the $m = 2, 3, 4$ colour-ordered vertices in dimensional regularisation, using the "X" Mathematica package [195] to perform the loop integrals.

B.1 $m = 2$

At two points in SDYM, the relevant diagram is the bubble, given by

$$\mathcal{V}_2^{(1)} = \mu^{2\epsilon} \int \frac{d^D l}{(2\pi)^D} \frac{X(l,k)X(l+k,-k)}{l^2(k+l^2)} = -\mu^{2\epsilon} \int \frac{d^D l}{(2\pi)^D} \frac{X(l,k)^2}{l^2(k+l^2)}. \tag{B.1}$$

Here k and l are the external and loop momenta respectively, and μ is the regularisation scale. The loop momenta is $(4-2\epsilon)$-dimensional. We now employ some standard techniques. Firstly, we introduce the Feynman parametrisation,

$$\frac{1}{l^2(k+l^2)} = \int_0^1 dx \frac{1}{[l^2 + x((k+1)-l^2)]^2} = \int_0^1 dx \frac{1}{[(l+xk)^2 - \Delta]^2}, \tag{B.2}$$

where $\Delta = -xk^2(1-x)$. Then, we perform the shift $l \to \tilde{l} = l + xk$, such that eq. (B.1) becomes

$$\mathcal{V}_2^{(1)} = -\mu^{2\epsilon} \int_0^1 dx \int \frac{d^D\tilde{l}}{(2\pi)^D} \frac{X(\tilde{l},k)^2}{(\tilde{l}^2 - \Delta)^2} \tag{B.3}$$

$$= -\langle\eta|\sigma^\mu k|\eta\rangle\langle\eta|\sigma^\nu k|\eta\rangle\mu^{2\epsilon} \int_0^1 dx \int \frac{d^D\tilde{l}}{(2\pi)^D} \frac{\tilde{l}_\mu \tilde{l}_\nu}{(\tilde{l}^2 - \Delta)^2}. \tag{B.4}$$

The denominator is symmetric in $\tilde{l}$, and thus we can make the replacement [35]

$$\tilde{l}_\mu \tilde{l}_\nu \to \frac{1}{D}\tilde{l}^2 \eta_{\mu\nu}. \tag{B.5}$$

The metric contracts with the indices in the prefactor, and by applying the identity $\sigma^\mu_{\alpha\dot{\alpha}}\sigma_{\mu\beta\dot{\beta}} = 2\epsilon_{\alpha\beta}\epsilon_{\dot{\alpha}\dot{\beta}}$, we find that

$$\langle\eta|\sigma^\mu k|\eta\rangle\langle\eta|\sigma_\mu k|\eta\rangle \propto X(k,k) = 0. \tag{B.6}$$

We therefore see that $\mathcal{V}_2^{(1)} = 0$, even off-shell. There is a subtle issue of regularisation (see e.g. [176, 177]), but our choice here is consistent with the following higher-point computations.

A similar calculation can be carried out in SDG, where the integrand is constructed following the double copy prescription discussed in [46]. We find, similarly, that $\mathcal{V}_2^{(1)} = 0$ off-shell.

B.2 $m = 3$

At three points in SDYM, the relevant diagram is the triangle, given by

$$\mathcal{V}_3^{(1)} = \mu^{2\epsilon} \int \frac{d^D l}{(2\pi)^D} \frac{X(l,1)X(l,2)X(l+2,3)}{l^2(l-k_1)^2(l+k_2)^2}, \tag{B.7}$$

Due to the linearity of $X(k,k')$ in its arguments this reduces to a sum of rank-3 and rank-2 tensor integrals. These integrals can be evaluated with Passarino-Veltman reductions using the “X” Mathematica package, with the result

$$\mathcal{V}_3^{(1)} = \frac{i}{16\pi^2} X(2,3)^3 \left[\sum_{i=1}^{3} a_i \ln\left(-\frac{4\pi\mu^2}{k_i^2}\right) + bS_{c_0}(k_1^2,k_2^2,k_3^2) + \mathcal{R} \right], \tag{B.8}$$

where the coefficients are

$$a_1 = -\frac{k_1^2}{3\lambda}\left[(k_2^2+k_3^2)(k_2^2+k_3^2-k_1^2)^3 + 4k_2^2k_3^2(k_2^4+k_3^4-k_1^4)\right.$$
$$\left. +18k_1^2k_2^2k_3^2(k_2^2+k_3^2-k_1^2) - 24k_2^3k_3^4\right], \tag{B.9}$$

$$a_2 = a_1|_{k_1 \leftrightarrow k_2}, \tag{B.10}$$

$$a_3 = a_1|_{k_1 \leftrightarrow k_3}, \tag{B.11}$$

$$b = \frac{2k_1^2k_2^2k_3^2}{\lambda^3}\left[k_1^4(k_2^2+k_3^2-k_1^2) + k_2^4(k_3^2+k_1^2-k_2^2) + k_3^4(k_1^2+k_2^2-k_3^2) + 4k_1^2k_2^2k_3^2\right]. \tag{B.12}$$

Here, $\lambda = \lambda(k_1^2, k_2^2, k_3^2)$ is the Källén function

$$\lambda = 2(k_1^4+k_2^4+k_3^4) - (k_1^2+k_2^2+k_3^2)^2. \tag{B.13}$$

Finally, we have the rational part

$$\mathcal{R} = \frac{1}{6\lambda^2}\left[k_1^4(k_2^2+k_3^2-k_1^2) + k_2^4(k_3^2+k_1^2-k_2^2) + k_3^4(k_1^2+k_2^2-k_3^2) + 14k_1^2k_2^2k_3^2\right], \tag{B.14}$$

and the scalar function

$$S_{c_0}(k_1^2,k_2^2,k_3^2) = \frac{1}{\sqrt{\lambda}}\sum_{\text{cyc}(1,2,3)}\left[\mathrm{Li}_2\left(\frac{k_1^2+k_2^2-k_3^2+\sqrt{\lambda}}{k_1^2+k_2^2-k_3^2-\sqrt{\lambda}}\right)\right.$$
$$\left. -\mathrm{Li}_2\left(\frac{k_1^2+k_2^2-k_3^2-\sqrt{\lambda}}{k_1^2+k_2^2-k_3^2+\sqrt{\lambda}}\right)\right], \tag{B.15}$$

where Li_2 is a dilogarithm and the sum is over cyclic permutations.

In the limit where two legs are taken on-shell, only the rational part survives and we are left with

$$\mathcal{V}_3^{(1)}\Big|_{k_2^2,k_3^2\to 0} = \frac{i}{96\pi^2}\frac{X(2,3)^3}{k_1^2}. \tag{B.16}$$

When the final leg is taken on-shell this vanishes due to 3-point massless kinematics. This is as expected, since the 3-point all-plus amplitude vanishes.

The expression in eq. (B.8) for the 3-point one-loop vertex is not particularly appealing. However, it turns out that when computing the contribution of this 3-point vertex at 4-points, only the rational part is required. Despite this, the one-loop 3-point amplitude itself vanishes, and it is therefore inconvenient to use a formalism in which the 3-point vertex is non-zero. This is one of the reasons that the formalism based on anomaly cancellation on twistor space is more desirable, as here the first effective vertex appears at 4-points.

The same computation can be done in SDG, where we square the numerator in eq. (B.7). This yields rank-four, -five, and -six tensor integrals, which can be evaluated in Mathematica. The result takes a similar form to the SDYM case in eq. (B.8), but with more complicated coefficients and an overall factor of $X(2,3)^6$. Taking two of the legs on-shell, only the rational part survives, and we obtain

$$\text{SDG}: \qquad \mathcal{V}_3^{(1)}\big|_{k_2^2,k_3^2\to 0} \propto \frac{X(2,3)^6}{k_1^2}, \tag{B.17}$$

up to a numerical factor. This vanishes when the final leg goes on-shell, as expected.

B.3 $m = 4$

At four points, the computation is significantly more involved. The relevant diagram is the off-shell box. Once the loop integral is evaluated, the result has a similar structure as at three points, containing a rational part, logarithms, and dilogarithms, albeit with a far greater number of these terms.

We will not write the resulting vertex here, but to illustrate a point we consider instead the "off-shell amplitude" (without external polarisation factors). This is given by the sum of the 4-point one-loop vertex and the four diagrams with the one-loop 3-point vertex attached to the tree-level 3-point vertex. Taking all external legs except the first on-shell, only the rational parts of the loop vertices contribute, and after some massaging we obtain the following expression

$$\begin{aligned} I_4^{(1)}\big|_{k_2^2,k_3^2,k_4^2\to 0} \propto{}& \frac{X(1,2)X(3,4)^3}{s_{34}^2} + \frac{X(2,3)^3X(4,1)}{s_{23}^2} \\ &+ \frac{X(2,3)X(3,4)}{s_{23}s_{34}}\left[X(1,2)X(2,3) + X(3,4)X(4,1) + X(1,2)X(4,1)\right]. \end{aligned} \tag{B.18}$$

When the final leg is taken on-shell, this collapses to

$$I_4^{(1)}\big|_{k_1^2,k_2^2,k_3^2,k_4^2\to 0} \propto \frac{X(1,2)X(2,3)X(3,4)X(4,1)}{s_{12}s_{23}}, \tag{B.19}$$

which leads directly to the correct 4-point amplitude. The fact that only the rational parts contribute in these limits means that at this multiplicity we may take the rational parts to be the vertices. It is possible that this extends to all multiplicity and that this procedure therefore leads to rational vertices at all orders. However, we have highlighted the disadvantages of this formalism in this appendix, and these deficiencies are only exacerbated in SDG.

It is interesting to note that eqs. (B.16) and (B.18), corresponding to the off-shell amplitudes with all but one legs on-shell, constitute the one-loop Berends-Giele currents in this approach. Berends-Giele currents are recursively constructed perturbative solutions to equations of motion [196]. The one-loop currents above then correspond to perturbative solutions for $\bar{\Psi}$ in the quantum-corrected theory. These solutions are generated by considering the Ψ fields in the equations of motion as on-shell sources, such that $\bar{\Psi}$ is generated by positive-helicity sources only. In such an approach, one takes the "measured" field $\bar{\Psi}$ to be outgoing, while the sources are incoming. Thus, when we take all legs to be incoming the one-loop currents correspond to all-plus diagrams. Unfortunately, we were unable to find a closed form solution to the quantum-corrected equation of motion in eq. (6.46). We hoped that this would have similar features to the known perturbative solution to the other equation of motion in eq. (6.45) (see [171,172]), whose form is unchanged in the quantum theory. Despite this, we note that the one-loop currents, even in this brute-force approach, are rational functions up to 4-points.

Bibliography

[1] L. Alfonsi, C. D. White, and S. Wikeley, *Topology and Wilson lines: global aspects of the double copy*, *JHEP* **07** (2020) 091, [arXiv:2004.07181].

[2] R. Alawadhi, D. S. Berman, C. D. White, and S. Wikeley, *The single copy of the gravitational holonomy*, *JHEP* **10** (2021) 229, [arXiv:2107.01114].

[3] K. Armstrong-Williams, C. D. White, and S. Wikeley, *Non-perturbative aspects of the self-dual double copy*, *JHEP* **08** (2022) 160, [arXiv:2205.02136].

[4] R. Monteiro, R. Stark-Muchão, and S. Wikeley, *Anomaly and double copy in quantum self-dual Yang-Mills and gravity*, arXiv:2211.12407.

[5] Z. Bern, J. J. Carrasco, M. Chiodaroli, H. Johansson, and R. Roiban, *The Duality Between Color and Kinematics and its Applications*, 9, 2019.

[6] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016), no. 6 061102, [arXiv:1602.03837].

[7] R. Britto, F. Cachazo, B. Feng, and E. Witten, *Direct proof of tree-level recursion relation in Yang-Mills theory*, *Phys. Rev. Lett.* **94** (2005) 181602, [hep-th/0501052].

[8] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *One loop n point gauge theory amplitudes, unitarity and collinear limits*, *Nucl. Phys. B* **425** (1994) 217–260, [hep-ph/9403226].

[9] Z. Bern, J. J. M. Carrasco, and H. Johansson, *New Relations for Gauge-Theory Amplitudes*, *Phys. Rev. D* **78** (2008) 085011, [arXiv:0805.3993].

[10] Z. Bern, J. J. M. Carrasco, and H. Johansson, *Perturbative Quantum Gravity as a Double Copy of Gauge Theory*, *Phys. Rev. Lett.* **105** (2010) 061602, [arXiv:1004.0476].

[11] N. E. J. Bjerrum-Bohr, P. H. Damgaard, T. Sondergaard, and P. Vanhove, *The Momentum Kernel of Gauge and Gravity Theories*, *JHEP* **01** (2011) 001, [arXiv:1010.3933].

[12] C. R. Mafra, O. Schlotterer, and S. Stieberger, *Explicit BCJ Numerators from Pure Spinors*, *JHEP* **07** (2011) 092, [arXiv:1104.5224].

[13] C.-H. Fu, Y.-J. Du, and B. Feng, *An algebraic approach to BCJ numerators*, *JHEP* **03** (2013) 050, [arXiv:1212.6168].

[14] N. E. J. Bjerrum-Bohr, J. L. Bourjaily, P. H. Damgaard, and B. Feng, *Manifesting Color-Kinematics Duality in the Scattering Equation Formalism*, *JHEP* **09** (2016) 094, [arXiv:1608.00006].

[15] L. de la Cruz, A. Kniss, and S. Weinzierl, *Properties of scattering forms and their relation to associahedra*, *JHEP* **03** (2018) 064, [arXiv:1711.07942].

[16] H. Elvang and Y.-t. Huang, *Scattering Amplitudes in Gauge Theory and Gravity.* Cambridge University Press, 4, 2015.

[17] M. Campiglia and S. Nagy, *A double copy for asymptotic symmetries in the self-dual sector*, *JHEP* **03** (2021) 262, [arXiv:2102.01680].

[18] R. Alawadhi, D. S. Berman, B. Spence, and D. Peinador Veiga, *S-duality and the double copy*, *JHEP* **03** (2020) 059, [arXiv:1911.06797].

[19] Y.-T. Huang, U. Kol, and D. O'Connell, *Double copy of electric-magnetic duality*, *Phys. Rev. D* **102** (2020), no. 4 046005, [arXiv:1911.06318].

[20] C. Cheung and J. Mangan, *Scattering Amplitudes and the Navier-Stokes Equation*, arXiv:2010.15970.

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

[22] A. Luna, R. Monteiro, I. Nicholson, and D. O'Connell, *Type D Spacetimes and the Weyl Double Copy*, *Class. Quant. Grav.* **36** (2019) 065003, [arXiv:1810.08183].

[23] C. D. White, *Twistorial Foundation for the Classical Double Copy*, *Phys. Rev. Lett.* **126** (2021), no. 6 061602, [arXiv:2012.02479].

[24] E. Chacón, S. Nagy, and C. D. White, *The Weyl double copy from twistor space*, *JHEP* **05** (2021) 2239, [arXiv:2103.16441].

[25] E. Chacón, S. Nagy, and C. D. White, *Alternative formulations of the twistor double copy*, *JHEP* **03** (2022) 180, [arXiv:2112.06764].

[26] T. Adamo, E. Casali, L. Mason, and S. Nekovar, *Scattering on plane waves and the double copy*, *Class. Quant. Grav.* **35** (2018), no. 1 015004, [arXiv:1706.08925].

[27] C. Armstrong, A. E. Lipstein, and J. Mei, *Color/kinematics duality in* AdS_4, *JHEP* **02** (2021) 194, [arXiv:2012.02059].

[28] A. Lipstein and S. Nagy, *Self-dual gravity and color/kinematics duality in* AdS_4, arXiv:2304.07141.

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

[30] C.-H. Fu and K. Krasnov, *Colour-Kinematics duality and the Drinfeld double of the Lie algebra of diffeomorphisms*, *JHEP* **01** (2017) 075, [arXiv:1603.02033].

[31] G. Chen, H. Johansson, F. Teng, and T. Wang, *On the kinematic algebra for BCJ numerators beyond the MHV sector*, *JHEP* **11** (2019) 055, [arXiv:1906.10683].

[32] G. Chen, H. Johansson, F. Teng, and T. Wang, *Next-to-MHV Yang-Mills kinematic algebra*, *JHEP* **10** (2021) 042, [arXiv:2104.12726].

[33] A. Brandhuber, G. Chen, H. Johansson, G. Travaglini, and C. Wen, *Kinematic Hopf Algebra for Bern-Carrasco-Johansson Numerators in Heavy-Mass Effective Field Theory and Yang-Mills Theory*, *Phys. Rev. Lett.* **128** (2022), no. 12 121601, [arXiv:2111.15649].

[34] G. Chen, G. Lin, and C. Wen, *Kinematic Hopf algebra for amplitudes and form factors*, arXiv:2208.05519.

[35] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory.* Addison-Wesley, Reading, USA, 1995.

[36] L. J. Dixon, *Calculating scattering amplitudes efficiently*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 95): QCD and Beyond*, pp. 539–584, 1, 1996. hep-ph/9601359.

[37] B. S. DeWitt, *Quantum Theory of Gravity. 3. Applications of the Covariant Theory*, *Phys. Rev.* **162** (1967) 1239–1256.

[38] J. M. Henn and J. C. Plefka, *Scattering Amplitudes in Gauge Theories*, vol. 883. Springer, Berlin, 2014.

[39] Z. Bern, T. Dennen, Y.-t. Huang, and M. Kiermaier, *Gravity as the Square of Gauge Theory*, *Phys. Rev. D* **82** (2010) 065003, [arXiv:1004.0693].

[40] N. E. J. Bjerrum-Bohr, P. H. Damgaard, R. Monteiro, and D. O'Connell, *Algebras for Amplitudes*, *JHEP* **06** (2012) 061, [arXiv:1203.0944].

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

[42] N. E. J. Bjerrum-Bohr, P. H. Damgaard, and P. Vanhove, *Minimal Basis for Gauge Theory Amplitudes*, *Phys. Rev. Lett.* **103** (2009) 161602, [arXiv:0907.1425].

[43] N. E. J. Bjerrum-Bohr, P. H. Damgaard, T. Sondergaard, and P. Vanhove, *Monodromy and Jacobi-like Relations for Color-Ordered Amplitudes*, *JHEP* **06** (2010) 003, [arXiv:1003.2403].

[44] S. Stieberger, *Open & Closed vs. Pure Open String Disk Amplitudes*, arXiv:0907.2211.

[45] B. Feng, R. Huang, and Y. Jia, *Gauge Amplitude Identities by On-shell Recursion Relation in S-matrix Program*, *Phys. Lett. B* **695** (2011) 350–353, [arXiv:1004.3417].

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

[47] Z. Bern, J. J. M. Carrasco, L. J. Dixon, H. Johansson, and R. Roiban, *Simplifying Multiloop Integrands and Ultraviolet Divergences of Gauge Theory and Gravity Amplitudes*, *Phys. Rev. D* **85** (2012) 105014, [arXiv:1201.5366].

[48] J. J. M. Carrasco and H. Johansson, *Five-Point Amplitudes in N=4 Super-Yang-Mills Theory and N=8 Supergravity*, *Phys. Rev. D* **85** (2012) 025006, [arXiv:1106.4711].

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

[50] M. Tolotti and S. Weinzierl, *Construction of an effective Yang-Mills Lagrangian with manifest BCJ duality*, *JHEP* **07** (2013) 111, [arXiv:1306.2975].

[51] M. Ben-Shahar, L. Garozzo, and H. Johansson, *Lagrangians Manifesting Color-Kinematics Duality in the NMHV Sector of Yang-Mills*, arXiv:2301.00233.

[52] P. Ferrero and D. Francia, *On the Lagrangian formulation of the double copy to cubic order*, *JHEP* **02** (2021) 213, [arXiv:2012.00713].

[53] L. Borsten and S. Nagy, *The pure BRST Einstein-Hilbert Lagrangian from the double-copy to cubic order*, *JHEP* **07** (2020) 093, [arXiv:2004.14945].

[54] L. Borsten, B. Jurčo, H. Kim, T. Macrelli, C. Saemann, and M. Wolf, *Becchi-Rouet-Stora-Tyutin-Lagrangian Double Copy of Yang-Mills Theory*, *Phys. Rev. Lett.* **126** (2021), no. 19 191601, [arXiv:2007.13803].

[55] L. Borsten, H. Kim, B. Jurčo, T. Macrelli, C. Saemann, and M. Wolf, *Double Copy from Homotopy Algebras*, *Fortsch. Phys.* **69** (2021), no. 8-9 2100075, [arXiv:2102.11390].

[56] A. Parkes, *A Cubic action for selfdual Yang-Mills*, *Phys. Lett. B* **286** (1992) 265–270, [hep-th/9203074].

[57] J. F. Plebanski, *Some solutions of complex Einstein equations*, *J. Math. Phys.* **16** (1975) 2395–2402.

[58] A. Luna, R. Monteiro, I. Nicholson, A. Ochirov, D. O'Connell, N. Westerberg, and C. D. White, *Perturbative spacetimes from Yang-Mills theory*, *JHEP* **04** (2017) 069, [arXiv:1611.07508].

[59] W. D. Goldberger and A. K. Ridgway, *Radiation and the classical double copy for color charges*, *Phys. Rev. D* **95** (2017), no. 12 125010, [arXiv:1611.03493].

[60] A. Luna, I. Nicholson, D. O'Connell, and C. D. White, *Inelastic Black Hole Scattering from Charged Scalar Amplitudes*, *JHEP* **03** (2018) 044, [arXiv:1711.03901].

[61] C.-H. Shen, *Gravitational Radiation from Color-Kinematics Duality*, *JHEP* **11** (2018) 162, [arXiv:1806.07388].

[62] Z. Bern, J. J. Carrasco, M. Chiodaroli, H. Johansson, and R. Roiban, *The SAGEX review on scattering amplitudes Chapter 2: An invitation to color-kinematics duality and the double copy*, *J. Phys. A* **55** (2022), no. 44 443003, [arXiv:2203.13013].

[63] D. A. Kosower, B. Maybee, and D. O'Connell, *Amplitudes, Observables, and Classical Scattering*, *JHEP* **02** (2019) 137, [arXiv:1811.10950].

[64] T. Adamo, A. Cristofoli, and A. Ilderton, *Classical physics from amplitudes on curved backgrounds*, *JHEP* **08** (2022) 281, [arXiv:2203.13785].

[65] T. Adamo, A. Cristofoli, A. Ilderton, and S. Klisch, *All-order waveforms from amplitudes*, arXiv:2210.04696.

[66] C. Shi and J. Plefka, *Classical double copy of worldline quantum field theory*, *Phys. Rev. D* **105** (2022), no. 2 026007, [arXiv:2109.10345].

[67] F. Comberiati and C. Shi, *Classical Double Copy of Spinning Worldline Quantum Field Theory*, arXiv:2212.13855.

[68] T. Adamo and A. Ilderton, *Classical and quantum double copy of back-reaction*, *JHEP* **09** (2020) 200, [arXiv:2005.05807].

[69] H. Stephani, D. Kramer, M. A. H. MacCallum, C. Hoenselaers, and E. Herlt, *Exact solutions of Einstein's field equations.* Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, Cambridge, 2003.

[70] A. Luna, R. Monteiro, D. O'Connell, and C. D. White, *The classical double copy for taub–nut spacetime*, *Physics Letters B* **750** (Nov, 2015) 272–277.

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

[72] N. Bahjat-Abbas, A. Luna, and C. D. White, *The Kerr-Schild double copy in curved spacetime*, *JHEP* **12** (2017) 004, [arXiv:1710.01953].

[73] K. Kim, K. Lee, R. Monteiro, I. Nicholson, and D. Peinador Veiga, *The Classical Double Copy of a Point Charge*, *JHEP* **02** (2020) 046, [arXiv:1912.02177].

[74] D. S. Berman, E. Chacón, A. Luna, and C. D. White, *The self-dual classical double copy, and the Eguchi-Hanson instanton*, *JHEP* **01** (2019) 107, [arXiv:1809.04063].

[75] K. Lee, *Kerr-Schild Double Field Theory and Classical Double Copy*, *JHEP* **10** (2018) 027, [arXiv:1807.08443].

[76] W. Cho and K. Lee, *Heterotic Kerr-Schild Double Field Theory and Classical Double Copy*, *JHEP* **07** (2019) 030, [arXiv:1904.11650].

[77] D. S. Berman, K. Kim, and K. Lee, *The classical double copy for M-theory from a Kerr-Schild ansatz for exceptional field theory*, *JHEP* **04** (2021) 071, [arXiv:2010.08255].

[78] G. Alkac, M. K. Gumus, and M. Tek, *The Kerr-Schild Double Copy in Lifshitz Spacetime*, *JHEP* **05** (2021) 214, [arXiv:2103.06986].

[79] R. Dempsey and P. Weck, *Compactifying the Kerr-Schild Double Copy*, `arXiv:2211.14327`.

[80] H. Godazgar, M. Godazgar, R. Monteiro, D. Peinador Veiga, and C. N. Pope, *Weyl Double Copy for Gravitational Waves*, *Phys. Rev. Lett.* **126** (2021), no. 10 101103, [`arXiv:2010.02925`].

[81] R. Monteiro, D. O'Connell, D. Peinador Veiga, and M. Sergola, *Classical solutions and their double copy in split signature*, *JHEP* **05** (2021) 268, [`arXiv:2012.11190`].

[82] R. Monteiro, S. Nagy, D. O'Connell, D. Peinador Veiga, and M. Sergola, *NS-NS spacetimes from amplitudes*, *JHEP* **06** (2022) 021, [`arXiv:2112.08336`].

[83] E. Chacón, A. Luna, and C. D. White, *Double copy of the multipole expansion*, *Phys. Rev. D* **106** (2022), no. 8 086020, [`arXiv:2108.07702`].

[84] A. Anastasiou, L. Borsten, M. J. Duff, L. J. Hughes, and S. Nagy, *Yang-Mills origin of gravitational symmetries*, *Phys. Rev. Lett.* **113** (2014), no. 23 231606, [`arXiv:1408.4434`].

[85] A. Anastasiou, L. Borsten, M. J. Duff, S. Nagy, and M. Zoccali, *Gravity as Gauge Theory Squared: A Ghost Story*, *Phys. Rev. Lett.* **121** (2018), no. 21 211601, [`arXiv:1807.02486`].

[86] G. Lopes Cardoso, G. Inverso, S. Nagy, and S. Nampuri, *Comments on the double copy construction for gravitational theories*, *PoS* **CORFU2017** (2018) 177, [`arXiv:1803.07670`].

[87] A. Luna, S. Nagy, and C. White, *The convolutional double copy: a case study with a point*, *JHEP* **09** (2020) 062, [`arXiv:2004.11254`].

[88] N. Bahjat-Abbas, R. Stark-Muchão, and C. D. White, *Monopoles, shockwaves and the classical double copy*, *JHEP* **04** (2020) 102, [`arXiv:2001.09918`].

[89] R. A. Brandt and F. Neri, *Stability Analysis for Singular Nonabelian Magnetic Monopoles*, *Nucl. Phys. B* **161** (1979) 253–282.

[90] R. A. Brandt and F. Neri, *Magnetic Monopoles in SU(n) Gauge Theories*, *Nucl. Phys. B* **186** (1981) 84–108.

[91] S. R. Coleman, *The Magnetic Monopole Fifty Years Later*, in *Les Houches Summer School of Theoretical Physics: Laser-Plasma Interactions*, pp. 461–552, 6, 1982.

[92] E. J. Weinberg, *Classical solutions in quantum field theory: Solitons and Instantons in High Energy Physics.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 9, 2012.

[93] T. T. Wu and C. N. Yang, *Concept of Nonintegrable Phase Factors and Global Formulation of Gauge Fields*, *Phys. Rev. D* **12** (1975) 3845–3857.

[94] **MoEDAL** Collaboration, B. Acharya et al., *Search for magnetic monopoles produced via the Schwinger mechanism*, *Nature* **602** (2022), no. 7895 63–67, [arXiv:2106.11933].

[95] A. H. Taub, *Empty space-times admitting a three parameter group of motions*, *Annals of Mathematics* **53** (1951), no. 3 472–490.

[96] E. Newman, L. Tamburino, and T. Unti, *Empty-space generalization of the schwarzschild metric*, *Journal of Mathematical Physics* **4** (1963), no. 7 915–923, [https://doi.org/10.1063/1.1704018].

[97] T. Ortin, *Gravity and Strings.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2nd ed. ed., 7, 2015.

[98] C. W. Misner, *The flatter regions of newman, unti, and tamburino's generalized schwarzschild space*, *Journal of Mathematical Physics* **4** (1963), no. 7 924–937, [https://doi.org/10.1063/1.1704019].

[99] J. F. Plebañski, *A class of solutions of einstein-maxwell equations*, *Annals of Physics* **90** (1975), no. 1 196 – 255.

[100] Z.-W. Chong, G. Gibbons, H. Lü, and C. Pope, *Separability and killing tensors in kerr–taub-nut–de sitter metrics in higher dimensions*, *Physics Letters B* **609** (Mar, 2005) 124–132.

[101] H. Georgi, *Lie Algebras In Particle Physics : from Isospin To Unified Theories.* Taylor & Francis, Boca Raton, 2000.

[102] P. Goddard, J. Nuyts, and D. I. Olive, *Gauge Theories and Magnetic Charge*, *Nucl. Phys. B* **125** (1977) 1–28.

[103] J. Fuchs and C. Schweigert, *Symmetries, Lie algebras and representations: A graduate course for physicists.* Cambridge University Press, 10, 2003.

[104] Y. Aharonov and D. Bohm, *Significance of electromagnetic potentials in the quantum theory*, *Phys. Rev.* **115** (Aug, 1959) 485–491.

[105] L. Landau and E. Lifschits, *The Classical Theory of Fields*, vol. Volume 2 of *Course of Theoretical Physics.* Pergamon Press, Oxford, 1975.

[106] P. A. M. Dirac, *Quantised singularities in the electromagnetic field,*, *Proc. Roy. Soc. Lond. A* **133** (1931), no. 821 60–72.

[107] J. S. Dowker and J. A. Roche, *The Gravitational analogues of magnetic monopoles*, *Proc. Phys. Soc.* **92** (1967) 1–8.

[108] S. Mandelstam, *Quantization of the gravitational field*, *Annals Phys.* **19** (1962) 25–66.

[109] G. Modanese, *Geodesic round trips by parallel transport in quantum gravity*, *Phys. Rev. D* **47** (1993) 502–509.

[110] G. Modanese, *Wilson loops in four-dimensional quantum gravity*, *Phys. Rev. D* **49** (1994) 6534–6542, [hep-th/9307148].

[111] H. W. Hamber and R. M. Williams, *Gravitational Wilson loop and large scale curvature*, *Phys. Rev. D* **76** (2007) 084008, [arXiv:0706.2342].

[112] A. Brandhuber, P. Heslop, A. Nasti, B. Spence, and G. Travaglini, *Four-point Amplitudes in N=8 Supergravity and Wilson Loops*, *Nucl. Phys. B* **807** (2009) 290–314, [arXiv:0805.2763].

[113] H. W. Hamber and R. M. Williams, *Gravitational Wilson Loop in Discrete Quantum Gravity*, *Phys. Rev. D* **81** (2010) 084048, [arXiv:0907.2652].

[114] H. W. Hamber and R. M. Williams, *Newtonian potential in quantum Regge gravity*, *Nucl. Phys. B* **435** (1995) 361–398, [hep-th/9406163].

[115] S. G. Naculich and H. J. Schnitzer, *Eikonal methods applied to gravitational scattering amplitudes*, *JHEP* **05** (2011) 087, [arXiv:1101.1524].

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

[117] D. J. Miller and C. D. White, *The Gravitational cusp anomalous dimension from AdS space*, *Phys. Rev. D* **85** (2012) 104034, [arXiv:1201.2358].

[118] S. Melville, S. G. Naculich, H. J. Schnitzer, and C. D. White, *Wilson line approach to gravity in the high energy limit*, *Phys. Rev. D* **89** (2014), no. 2 025009, [arXiv:1306.6019].

[119] A. Luna, S. Melville, S. G. Naculich, and C. D. White, *Next-to-soft corrections to high energy scattering in QCD and gravity*, *JHEP* **01** (2017) 052, [arXiv:1611.02172].

[120] S. Oxburgh and C. D. White, *BCJ duality and the double copy in the soft limit*, *JHEP* **02** (2013) 127, [arXiv:1210.1110].

[121] A. Sabio Vera, E. Serna Campillo, and M. A. Vazquez-Mozo, *Color-Kinematics Duality and the Regge Limit of Inelastic Amplitudes*, *JHEP* **04** (2013) 086, [arXiv:1212.5103].

[122] H. Johansson, A. Sabio Vera, E. Serna Campillo, and M. A. Vazquez-Mozo, *Color-kinematics duality and dimensional reduction for graviton emission in Regge limit*, in *International Workshop on Low X Physics*, 10, 2013. arXiv:1310.1680.

[123] M. Berger, *Sur les groupes d'holonomie homogènes de variétés à connexion affine et des variétés riemanniennes*, *Bulletin de la Société Mathématique de France* **83** (1955) 279–330.

[124] M. Berger, *Sur les groupes d'holonomie homogènes de variétés à connexion affine et des variétés riemanniennes*, *Bulletin de la Société Mathématique de France* **83** (1955) 279–330.

[125] A. L. Besse, *Einstein Manifolds*. Springer-Verlag, Berlin, Heidelberg, New York, 1987.

[126] A. Clarke and B. Santoro, *Holonomy groups in riemannian geometry*, 2012.

[127] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 12, 2007.

[128] D. R. Green, *Worldlines as Wilson Lines*, *Phys. Rev. D* **78** (2008) 064066, [arXiv:0804.4450].

[129] E. Gardi and L. Magnea, *Infrared singularities in QCD amplitudes*, *Nuovo Cim. C* **32N5-6** (2009) 137–157, [arXiv:0908.3273].

[130] C. D. White, *An Introduction to Webs*, *J. Phys. G* **43** (2016), no. 3 033002, [arXiv:1507.02167].

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

[132] E. Laenen, G. Stavenga, and C. D. White, *Path integral approach to eikonal and next-to-eikonal exponentiation*, *JHEP* **03** (2009) 054, [arXiv:0811.2067].

[133] M. Levi, *Effective Field Theories of Post-Newtonian Gravity: A comprehensive review*, *Rept. Prog. Phys.* **83** (2020), no. 7 075901, [arXiv:1807.01699].

[134] J. S. Dowker and J. A. Roche, *The gravitational analogues of magnetic monopoles*, *Proceedings of the Physical Society* **92** (sep, 1967) 1–8.

[135] J. Li and S. G. Prabhu, *Gravitational radiation from the classical spinning double copy*, *Phys. Rev. D* **97** (2018), no. 10 105019, [arXiv:1803.02405].

[136] W. D. Goldberger, S. G. Prabhu, and J. O. Thompson, *Classical gluon and graviton radiation from the bi-adjoint scalar double copy*, *Phys. Rev. D* **96** (2017), no. 6 065009, [arXiv:1705.09263].

[137] W. D. Goldberger, J. Li, and S. G. Prabhu, *Spinning particles, axion radiation, and the classical double copy*, *Phys. Rev. D* **97** (2018), no. 10 105018, [arXiv:1712.09250].

[138] F. Cachazo and A. Strominger, *Evidence for a New Soft Graviton Theorem*, arXiv:1404.4091.

[139] E. Casali, *Soft sub-leading divergences in Yang-Mills amplitudes*, *JHEP* **08** (2014) 077, [arXiv:1404.5551].

[140] C. D. White, *Diagrammatic insights into next-to-soft corrections*, *Phys. Lett. B* **737** (2014) 216–222, [arXiv:1406.7184].

[141] A. Sabio Vera and M. A. Vazquez-Mozo, *The Double Copy Structure of Soft Gravitons*, *JHEP* **03** (2015) 070, [arXiv:1412.3699].

[142] T. Rothman, G. F. R. Ellis, and J. Murugan, *Holonomy in the Schwarzschild-Droste geometry*, *Class. Quant. Grav.* **18** (2001) 1217–1234, [gr-qc/0008070].

[143] D. Bini, C. Cherubini, and R. T. Jantzen, *Circular holonomy in the Taub NUT space-time*, *Class. Quant. Grav.* **19** (2002) 5481–5488, [gr-qc/0210003].

[144] C. D. White, *Exact solutions for the biadjoint scalar field*, *Phys. Lett. B* **763** (2016) 365–369, [arXiv:1606.04724].

[145] P.-J. De Smet and C. D. White, *Extended solutions for the biadjoint scalar field*, *Phys. Lett. B* **775** (2017) 163–167, [arXiv:1708.01103].

[146] N. Bahjat-Abbas, R. Stark-Muchão, and C. D. White, *Biadjoint wires*, *Phys. Lett. B* **788** (2019) 274–279, [arXiv:1810.08118].

[147] C. Cheung, J. Mangan, J. Parra-Martinez, and N. Shah, *Non-perturbative Double Copy in Flatland*, *Phys. Rev. Lett.* **129** (2022), no. 22 221602, [arXiv:2204.07130].

[148] S. R. Coleman, *The Uses of Instantons*, *Subnucl. Ser.* **15** (1979) 805.

[149] N. Dorey, T. J. Hollowood, V. V. Khoze, and M. P. Mattis, *The Calculus of many instantons*, *Phys. Rept.* **371** (2002) 231–459, [hep-th/0206063].

[150] S. W. Hawking, *Gravitational Instantons*, *Phys. Lett. A* **60** (1977) 81.

[151] G. W. Gibbons and S. W. Hawking, *Gravitational Multi - Instantons*, *Phys. Lett. B* **78** (1978) 430.

[152] T. Eguchi and A. J. Hanson, *Selfdual Solutions to Euclidean Gravity*, *Annals Phys.* **120** (1979) 82.

[153] G. W. Gibbons and S. W. Hawking, *Classification of Gravitational Instanton Symmetries*, *Commun. Math. Phys.* **66** (1979) 291–310.

[154] T. Eguchi and A. J. Hanson, *GRAVITATIONAL INSTANTONS*, *Gen. Rel. Grav.* **11** (1979) 315–320.

[155] A. Luna, N. Moynihan, and C. D. White, *Why is the Weyl double copy local in position space?*, *JHEP* **12** (2022) 046, [arXiv:2208.08548].

[156] G. H. Derrick, *Comments on nonlinear wave equations as models for elementary particles*, *J. Math. Phys.* **5** (1964) 1252–1254.

[157] T. T. Wu and C.-N. Yang, *SOME SOLUTIONS OF THE CLASSICAL ISOTOPIC GAUGE FIELD EQUATIONS*, .

[158] G. 't Hooft, *Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle*, *Phys. Rev. D* **14** (1976) 3432–3450. [Erratum: Phys.Rev.D 18, 2199 (1978)].

[159] S. Vandoren and P. van Nieuwenhuizen, *Lectures on instantons*, arXiv:0802.1862.

[160] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld, and Y. I. Manin, *Construction of Instantons*, *Phys. Lett. A* **65** (1978) 185–187.

[161] M. K. Prasad, *Equivalence of Eguchi-Hanson metric to two-center Gibbons-Hawking metric*, *Phys. Lett. B* **83** (1979) 310–310.

[162] V. E. Zakharov and A. V. Mikhailov, *Relativistically Invariant Two-Dimensional Models in Field Theory Integrable by the Inverse Problem Technique. (In Russian)*, *Sov. Phys. JETP* **47** (1978) 1017–1027.

[163] C. Cheung, K. Kampf, J. Novotny, and J. Trnka, *Effective Field Theories from Soft Limits of Scattering Amplitudes*, *Phys. Rev. Lett.* **114** (2015), no. 22 221602, [arXiv:1412.4095].

[164] C. Cheung, K. Kampf, J. Novotny, C.-H. Shen, and J. Trnka, *On-Shell Recursion Relations for Effective Field Theories*, *Phys. Rev. Lett.* **116** (2016), no. 4 041601, [arXiv:1509.03309].

[165] K. Hinterbichler and A. Joyce, *Hidden symmetry of the Galileon*, *Phys. Rev. D* **92** (2015), no. 2 023503, [arXiv:1501.07600].

[166] J. Hoppe, *Diffeomorphism Groups, Quantization and SU(infinity)*, *Int. J. Mod. Phys. A* **4** (1989) 5235.

[167] C. Nash and S. Sen, *TOPOLOGY AND GEOMETRY FOR PHYSICISTS*. 1983.

[168] J. J. Oh, C. Park, and H. S. Yang, *Yang-Mills Instantons from Gravitational Instantons*, *JHEP* **04** (2011) 087, [arXiv:1101.1357].

[169] Z. Bern, G. Chalmers, L. J. Dixon, and D. A. Kosower, *One loop N gluon amplitudes with maximal helicity violation via collinear limits*, *Phys. Rev. Lett.* **72** (1994) 2134–2137, [hep-ph/9312333].

[170] Z. Bern, L. J. Dixon, M. Perelstein, and J. S. Rozowsky, *One loop n point helicity amplitudes in (selfdual) gravity*, *Phys. Lett. B* **444** (1998) 273–283, [hep-th/9809160].

[171] W. A. Bardeen, *Selfdual Yang-Mills theory, integrability and multiparton amplitudes*, *Prog. Theor. Phys. Suppl.* **123** (1996) 1–8.

[172] D. Cangemi, *Selfdual Yang-Mills theory and one loop like - helicity QCD multi - gluon amplitudes*, *Nucl. Phys. B* **484** (1997) 521–537, [hep-th/9605208].

[173] A. A. Rosly and K. G. Selivanov, *On amplitudes in selfdual sector of Yang-Mills theory*, *Phys. Lett. B* **399** (1997) 135–140, [hep-th/9611101].

[174] A. Brandhuber, B. Spence, and G. Travaglini, *Amplitudes in Pure Yang-Mills and MHV Diagrams*, *JHEP* **02** (2007) 088, [hep-th/0612007].

[175] R. Boels and C. Schwinn, *Deriving CSW rules for massive scalar legs and pure Yang-Mills loops*, *JHEP* **07** (2008) 007, [arXiv:0805.1197].

[176] P. Chattopadhyay and K. Krasnov, *One-loop same helicity four-point amplitude from shifts*, *JHEP* **06** (2020) 082, [arXiv:2002.11390].

[177] P. Chattopadhyay and K. Krasnov, *One-loop same helicity YM amplitudes from BG currents*, *JHEP* **03** (2022) 191, [arXiv:2110.00331].

[178] K. J. Costello, *Quantizing local holomorphic field theories on twistor space*, arXiv:2111.08879.

[179] K. Costello and N. M. Paquette, *Celestial holography meets twisted holography: 4d amplitudes from chiral correlators*, *JHEP* **10** (2022) 193, [arXiv:2201.02595].

[180] R. Bittleston, A. Sharma, and D. Skinner, *Quantizing the non-linear graviton*, arXiv:2208.12701.

[181] K. Lee, *Quantum off-shell recursion relation*, *JHEP* **05** (2022) 051, [arXiv:2202.08133].

[182] H. Gomez, R. Lipinski Jusinskas, C. Lopez-Arcos, and A. Quintero Velez, *One-Loop Off-Shell Amplitudes from Classical Equations of Motion*, *Phys. Rev. Lett.* **130** (2023), no. 8 081601, [arXiv:2208.02831].

[183] H. Kakkad, P. Kotko, and A. Stasto, *One-Loop effective action approach to quantum MHV theory*, *JHEP* **11** (2022) 132, [arXiv:2208.11000].

[184] G. Chalmers and W. Siegel, *The Selfdual sector of QCD amplitudes*, *Phys. Rev. D* **54** (1996) 7628–7633, [hep-th/9606061].

[185] G. Chalmers and W. Siegel, *Simplifying algebra in Feynman graphs. Part 2. Spinor helicity from the space-cone*, *Phys. Rev. D* **59** (1999) 045013, [hep-ph/9801220].

[186] W. Siegel, *Selfdual N=8 supergravity as closed N=2 (N=4) strings*, *Phys. Rev. D* **47** (1993) 2504–2511, [hep-th/9207043].

[187] G. Mahlon, *One loop multi - photon helicity amplitudes*, *Phys. Rev. D* **49** (1994) 2197–2210, [hep-ph/9311213].

[188] M. Dunajski, *Solitons, instantons, and twistors.* 2010.

[189] E. Chacón, H. García-Compeán, A. Luna, R. Monteiro, and C. D. White, *New heavenly double copies*, *JHEP* **03** (2021) 247, [arXiv:2008.09603].

[190] A. Brandhuber, B. Spence, G. Travaglini, and K. Zoubos, *One-loop MHV Rules and Pure Yang-Mills*, *JHEP* **07** (2007) 002, [arXiv:0704.0245].

[191] M. B. Green and J. H. Schwarz, *Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory*, *Phys. Lett. B* **149** (1984) 117–122.

[192] L. Boyle and N. Turok, *Cancelling the vacuum energy and Weyl anomaly in the standard model with dimension-zero scalar fields*, arXiv:2110.06258.

[193] Z. Bern, H.-H. Chi, L. Dixon, and A. Edison, *Two-Loop Renormalization of Quantum Gravity Simplified*, *Phys. Rev. D* **95** (2017), no. 4 046013, [arXiv:1701.02422].

[194] A. Chakrabarti and D. H. Tchrakian, *Antisymmetrized 2p forms generalizing curvature two forms and a corresponding p hierarchy of Schwarzschild type metrics in dimensions $d > 2p + 1$*, *Adv. Theor. Math. Phys.* **3** (1999) 791–814, [hep-th/9908128].

[195] H. H. Patel, *Package-X 2.0: A Mathematica package for the analytic calculation of one-loop integrals*, *Comput. Phys. Commun.* **218** (2017) 66–70, [arXiv:1612.00009].

[196] F. A. Berends and W. T. Giele, *Recursive Calculations for Processes with n Gluons*, *Nucl. Phys. B* **306** (1988) 759–808.