

# **Gravitational Waves from Feynman Diagrams**

Lucien Huber

9/13/2022

The goal of this thesis is to explain the nascent field of scattering amplitudes in gravitational wave physics. We explore the context in which these tools are revolutionizing the field, and then go into detail of their construction and application.

We begin with a brief overview of the topic and the academic context. The next chapter is devoted to the theoretical foundations of gravitational waves, and their important sources. In the third chapter we discuss the detection of these waves, highlighting the importance of the theoretical templates. The fourth chapter explores the theoretical methods for producing full wave templates, such as Effective One-Body (EOB) and Numerical Relativity (NR) as well as the expansions and approximations used for the inspiral: the Post-Newtonian (PN) and Post-Minkowski (PM) approximations.

The fifth chapter is the main focus of the thesis, and is devoted to the scattering amplitude techniques that can be applied to the general relativistic two-body problem. We begin with the Effective Field Theory (EFT) matching technique to motivate the production of gravitational amplitudes. The Kosower Maybee and O'Connell (KMOC) formalism, defining the classical limit of wave functions, is explained in detail. Finally, amplitudes in a simplified theory, Scalar Quantum Electrodynamics (SQED) are computed, with a focus on automation of calculations. The corresponding graphs, filters, Feynman rules and further simplifications are all implemented in a cocktail of the relevant languages. We also compare the results obtained with those found in the literature.

In the last chapter we provide some concluding remarks and a brief outlook.

# Table of contents

<b>Preface</b>	<b>5</b>
<b>Acknowledgements</b>	<b>6</b>
<b>1 Introduction</b>	<b>7</b>
<b>2 Gravitational Wave Generation</b>	<b>10</b>
2.1 Gravitational Waves (GWs) . . . . .	10
2.2 Homogeneous solutions . . . . .	13
2.3 Inhomogeneous solutions . . . . .	15
2.4 Gravitational Wave Sources . . . . .	17
2.4.1 Compact binaries . . . . .	20
2.4.2 Black Hole (BH) Binaries . . . . .	23
<b>3 Gravitational Wave Detection</b>	<b>25</b>
3.1 Laser interferometers . . . . .	25
3.1.1 Laser Interferometer Gravitational-Wave Observatory (LIGO) and ground based observatories . . . . .	26
3.1.2 Laser Interferometer Space Antenna (LISA) . . . . .	28
3.2 Matched filtering . . . . .	29
3.3 Pulsar Timing Array (PTA) . . . . .	31
<b>4 Waveform Generation</b>	<b>33</b>
4.1 Numerical Relativity . . . . .	34
4.2 Inspiral methods . . . . .	36
4.2.1 Post-Newtonian . . . . .	37
4.2.2 Post-Minkowski . . . . .	38
4.3 Effective One-Body . . . . .	40
<b>5 Scattering amplitudes and GWs</b>	<b>46</b>
5.1 Scattering amplitudes . . . . .	47
5.2 From amplitude to potential . . . . .	48
5.2.1 EFT amplitude . . . . .	50
5.3 KMOC framework . . . . .	54
5.3.1 Conventions . . . . .	54
5.3.2 Intital State . . . . .	55

5.3.3	Change in observable . . . . .	58
5.3.4	Classical limit . . . . .	60
5.4	Impulse in KMOC . . . . .	67
5.5	SQED amplitudes . . . . .	70
5.5.1	Expansions and simplifications . . . . .	73
5.5.2	Tree level . . . . .	75
5.5.3	One loop . . . . .	76
5.5.4	Higher order and gravity . . . . .	89
<b>6</b>	<b>Conclusion</b>	<b>93</b>
<b>References</b>		<b>95</b>

# Preface

This document represents my, Lucien Huber, master thesis, the last step in my master's degree at ETH Zurich. It can be viewed as either a webpage at <https://lcnhb.github.io/GWAmplitudes>, or downloaded as a PDF file. All code used for the thesis is provided in the GitHub repository <https://github.com/lcnhb/GWAmplitudeS>.

# Acknowledgements

First and foremost I would like to thank my supervisor, Prof. Babis Anastasiou, for his trust, his support, and kindness throughout the last half year that has led to this thesis. Without his curiosity towards this new subject, I would have never discovered and learned about this wonderful world at the confluence of two great physical theories, General Relativity (GR) and Quantum Field Theory (QFT). I would also like to extend this thanks to the rest of the theoretical particle physics group at ETH. Dr Achilleas Lazopoulos, Dr. Ben Ruijl, Dr. Valentin Hirschi, Rayan Haindl, Dario Kermanschah, Zeno Capatti, Matilde Vicini, and Julia Karlen, the many shared lunches, coffee breaks, and discussions have been an incredible source of inspiration and motivation for me. In particular, I would like to thank Ben for his help with all things FORM, Zeno for our very insightful discussions and your great suggestions, and Julia for reading through the first draft of this thesis.

I would also like to thank my girlfriend, Maria Miklos, for her patience, love, and support. Thanks for keeping me grounded and reminding me that there is more to life than physics (such as math)!

Thanks as well to my grandfather Martin Huber, who graciously hosted me during the start of my masters, and with which I shared many great discussions and stories at the breakfast, dinner, and Jass table.

Finally, I would like to thank my parents for their unconditional love and support, and without whom I would not be here today.

# 1 Introduction

The detection of GWs by the LIGO and Virgo collaborations in 2016 ([B. P. Abbott et al. 2016](#)) has sparked a new era of GW astronomy. The first detections were of Binary Black Hole (BBH) mergers. More recently, Binary Neutron Star (BNS) mergers ([B. P. Abbott et al. 2017](#)) as well as Neutron Star - Black Hole (NS-BH) mergers have been detected ([R. Abbott et al. 2021](#)). Future observatories will further increase sensitivity and will be able to detect a wide range of astrophysical sources. Studying these GWs signals gives us a very powerful new window into the universe. It allows us to study the properties of neutron stars and black holes, and the physics of compact object mergers, but also gives us a powerful testing apparatus for general relativity.

To detect these faint signals LIGO and Virgo have been made to be extraordinarily sensitive instruments. This sensitivity had been the main barrier to overcome ever since GW were first predicted by Einstein. Whilst sensitive measurements generally demand, and thus drive equally precise theoretical predictions, for GW astronomy, the case for theoretical models is even stronger. Indeed, theoretical models are necessary components to the detection mechanisms used at current laser interferometers. The faint signals of strain due to the passage of a GW are often buried in the noise of the detector. To counteract this, experimental physicists make use of matched filtering, an approach where a template signal is cross correlated to the detector output. The template is a model of the signal provided by theoretical physicists based on physical models. The upper bound to physical content of the detection is thus in fact set by the physical content of the template and therefore of the theoretical model. Additionally, the more precise the template, the higher the signal-to-noise ratio, the more probable and precise the detection. Theoretical models are thus a crucial component of the detection process, and there is a strong incentive to improve their precision and breadth of applicability.

In recent years, an unlikely ally in this precision race has arisen, based on QFT techniques ([Goldberger and Rothstein 2006](#),

[Kosower, Maybee, and O'Connell 2019](#); [Porto 2016](#); [Neill and Rothstein 2013](#)). This is seemingly an impossibility as QFT and GR have historically been the two irreconcilable theories that ‘together’ span the whole range of scales in the universe. QFT and the standard model, describe physics at the smallest scales, where quantum effects dominate. This is the physics of colliders, and condensed matter. The standard model is our best guess as to the nature of the fundamental constituents of ‘matter’ and ‘light’.

GR on the other hand describes physics at the largest scales, where gravity dominates. It was developed by Einstein as the successor to the Newtonian theory of gravity, which was not able to account for orbital mechanics such as the precession of the perihelion of mercury. Having passed every experimental test we have thrown at it, GR is by far the best theory of gravity we have yet. It can describe cosmological phenomena, as well as the compact objects known as black holes.

These two physics leviathans do not get along, unfortunately. GR is a theory of geometry, and QFT is a theory of interactions. We can try to write and interpret one in the language of the other, but in either case difficulties arise. Writing GR in terms of perturbative QFT one introduces the graviton, the spin 2 quantum of gravity. Disappointingly, such a theory breaks down at high energies, and is un-renormalisable . In fact the presence of black holes, that dominate the high energy spectrum of GR, is a testament to that fact ([Shomer 2007](#)). On the other hand, whilst one can write QFTs on a curved background, this does not quantize gravity itself, and therefore does not really address the problem. Background independent theories such as loop quantum gravity does not yet have semiclassical limit collapsing to GR. String theory, on the other hand, introduces spurious dimensions that need to be compactified, which leads to uncountably many solutions. In short, there is no satisfactory theory of that combines GR and QFT consistently.

Surprisingly, one can still use the powerful tools of one theory to tackle problems in the other. Specifically, considering the classical limit of QFT, allows the extraction of perturbative gravitational dynamics from an initially quantum description of the setup. This is interesting because the tools developed for high energy physics have become extremely powerful and efficient theoretical machines, often making use of diagrammatic reasoning. Tools such as Feynman diagrams, differential equations ([Remiddi 1997](#); [Zvi Bern, Dixon, and Kosower 1994](#); [Kotikov 1991](#)) and more recently generalized unitarity ([Z. Bern et al. 1995](#); [Zvi Bern et al.](#)

2005; Buchbinder and Cachazo 2005; Charalampos Anastasiou et al. 2007), all yielding efficient computations for scattering amplitudes, can all be applied to *classical* orbital dynamics in GR. This is exactly the subject of this thesis.

Specifically we are interested in the diagrammatic objects that arise when framing the two body problem similarly to particle collisions. Diagrammatics are interesting as they are an abstraction of the problem that enables efficient reasoning about and automation of the solutions. We will explain where these tools shine in the broader context of waveform approaches such as EOB , Nonrelativistic General Relativity (NRGR) , PM and PN approximations.

This thesis is organized as follows. In Chapter 2 we deal with the theoretical foundations of GWs, exploring how they arise through the linearized Einsteins' field equations. The subsequent Chapter 3 discusses the experimental apparatuses that are used to detect GWs. We describe at the techniques used to extract the signals from the noise, and motivate precise theoretical predictions. In Chapter 4 we look at how to generate partial and full waveforms. We explore the EOB and NR frameworks. Motivated by the need for more precise ingredients to input into EOB, in Chapter 5 we explain the state-of-the-art formalisms that incorporate amplitude tools into the program. We describe the conservative EFT matching approach, where the system does not lose energy to its environment. Taking into account the loss of energy due to GWs necessitates another method: the KMOC formalism, which makes up the second half of the chapter. In this context, we take a deeper look at amplitude computing tools. We conclude with a discussion of the challenges that lie ahead, and the future of GW astronomy.

## 2 Gravitational Wave Generation

We start by looking at where and how GWs are generated. We will see that GR predicts the existence of GWs, and we will explore under what conditions they are observable.

### 2.1 GWs

To this day, the most fundamental theory of gravity is GR, due to Einstein. It formulates spacetime as a Riemannian manifold with curvature, determined by mass distribution. This curvature is encoded in the metric  $g_{\mu\nu}$ . Objects then move in this deformed spacetime following Equation of Motions (EOMs) written in terms of  $g_{\mu\nu}$ . Interaction is thus caused by the dependence of  $g_{\mu\nu}$  on mass distribution, which itself evolves through  $g_{\mu\nu}$ . The physical laws that governs the interaction between mass and curvature are manifested in Einstein's field equations,

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2}R + \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu}. \quad (2.1)$$

The Left-Hand Side (LHS) of this equation contains several objects that only depend on the metric  $g_{\mu\nu}$ .  $R_{\mu\nu}$  is the Ricci tensor<sup>1</sup>, a contraction of the Riemann curvature tensor  $R^{\beta}_{\mu\nu\rho}$ . The Riemann tensor<sup>2</sup> encodes the curvature of spacetime, in terms of the non-commutativity of covariant derivatives. When they do, it means that they have collapsed to regular derivatives, and thus the Levi-civita connection<sup>3</sup> must have vanished. This is only possible if the metric is flat,  $g_{\mu\nu} = \eta_{\mu\nu}$ . The Riemann curvature tensor is exclusively composed of the metric, its first, second derivatives, and it is linear in the second derivative of the metric. In fact, it is the only possible tensor of this form. The second term on the LHS is the Ricci scalar, a further contraction of the Ricci Tensor:  $R = g^{\mu\nu}R_{\mu\nu}$ . The final term on the LHS is just a proportional

<sup>1</sup> The Ricci Tensor is given by:

$$R_{\mu\nu} := R^{\alpha}_{\mu\alpha\nu} = g^{\alpha\rho}R^{\alpha\mu\rho\nu} \quad (2.2)$$

<sup>2</sup> The Riemann Tensor is given by:

$$R^{\beta}_{\mu\nu\rho} = \Gamma^{\beta}_{\mu\nu,\rho} - \Gamma^{\beta}_{\mu\rho,\nu} - \Gamma^{\alpha}_{\mu\rho}\Gamma^{\beta}_{\alpha\nu} + \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\rho}. \quad (2.3)$$

<sup>3</sup> The Levi-Civita Connection is given by:

$$\Gamma^{\alpha}_{\mu\nu} = \frac{1}{2}g^{\alpha\rho}\{g_{\rho\nu,\mu} + g_{\rho\mu,\nu} - g_{\mu\nu,\rho}\} \quad (2.4)$$

constant factor of the metric,  $\Lambda$ , called the cosmological constant. It can be measured and has a low known upper bound. It is in part responsible for the expansion of the universe.

The Right-Hand Side (RHS) of eq. 2.1 encodes the distribution of mass and energy in the universe, through the stress-energy tensor  $T_{\mu\nu}$ , which depends on the dynamical properties of the system. It is further multiplied by the Gravitational constant  $G$ . In empty space,  $T_{\mu\nu}$  is zero.

Equation 2.1 can be recast in a form where the Ricci scalar has been substituted by the trace of the stress-energy tensor,

$$R_{\mu\nu} = -8\pi G \left( T_{\mu\nu} - T^\alpha_\alpha \frac{g_{\mu\nu}}{2} \right) + \Lambda g_{\mu\nu} \quad (2.5)$$

Both eq. 2.1 & eq. 2.5 with  $\Lambda = 0$  admit wave solutions. In order to see this, let us look at these equations in the weak field approximation. We take the metric to be a small perturbation around the Minkowskian background,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad (2.6)$$

where  $h_{\mu\nu}$  is the expansion parameter (thus considered small). Note that if we restrict ourselves to first order in  $h_{\mu\nu}$  then all raising and lowering of indices is done with  $\eta_{\mu\nu}$ . Since the LHS of eq. 2.1 is composed of the Ricci tensor and scalar, let us see how these behave in the weak approximation. Both are in fact made up of Levi-Civita connections, which to first order in  $h_{\mu\nu}$  is given by:

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2} \eta^{\alpha\rho} \{ h_{\rho\nu,\mu} + h_{\rho\mu,\nu} - h_{\mu\nu,\rho} \} + \mathcal{O}(h^2) . \quad (2.7)$$

Plugging eq. 2.7 into the definition of the Ricci tensor we see that the terms with products of the connection vanish, and we are left with the derivative terms:

$$R_{\mu\nu} = \frac{\eta^{\alpha\rho}}{2} [ h_{\rho\alpha,\mu\nu} + h_{\mu\nu,\rho\alpha} - h_{\mu\alpha,\rho\nu} - h_{\nu\rho,\mu\alpha} ] + \mathcal{O}(h^2) = R^{(1)}_{\mu\nu} + \mathcal{O}(h^2) \quad (2.8)$$

The linearized Ricci scalar is then just  $R^{(1)} = \eta^{\mu\nu} R^{(1)}_{\mu\nu}$ . The symmetry of the Einstein Field Equations (EFE) under general diffeomorphisms implies the existence of gauge freedom. In particular,

eq. 2.1, admits an infinite number of solutions, and in fact we can map a solution into another by changing coordinates in such a way that the equation isn't affected. To fix this ambiguity we choose a coordinate system ('gauge choice') by imposing the harmonic coordinate conditions:

$$g^{\alpha\beta}\Gamma_{\alpha\beta}^\mu = 0 \quad (2.9)$$

The harmonic coordinate conditions demand the vanishing of the trace of the Levi-Civita connection. They have a simplified form in the weak approximation, which we can access by plugging eq. 2.7 into eq. 2.9, giving:

$$\begin{aligned} (\eta^{\alpha\beta} + h^{\alpha\beta}) \frac{1}{2} (\eta^{\mu\rho} + h^{\mu\rho}) [h_{\alpha\rho,\beta} + h_{\beta\rho,\alpha} - h_{\alpha\beta,\rho}] &= 0 \\ \eta^{\mu\rho} \eta^{\alpha\beta} [2h_{\alpha\rho,\beta} - h_{\alpha\beta,\rho}] + \mathcal{O}(h^2) &= 0 \\ \eta^{\alpha\beta} h_{\alpha\rho,\beta} &= \frac{1}{2} h_{\alpha\beta,\rho} \eta^{\alpha\beta} + \mathcal{O}(h^2). \end{aligned} \quad (2.10)$$

The last equation, truncated to first order, is called the De-Donder gauge condition:

$$h_{\mu\nu}{}^\mu = \frac{1}{2} h_{\alpha,\nu}^\alpha \quad (2.11)$$

In the de-Donder gauge we can simplify the terms in eq. 2.8 to:

$$\eta^{\alpha\beta} h_{\alpha\mu,\nu\beta} = \frac{1}{2} \eta^{\alpha\beta} h_{\alpha\beta,\mu\nu}, \quad (2.12)$$

and

$$\eta^{\alpha\beta} h_{\beta\nu,\mu\alpha} = \frac{1}{2} h_{\alpha\beta,\mu\nu}. \quad (2.13)$$

With these two relations the expression for the linearized Ricci tensor eq. 2.8 simplifies to

$$R^{(1)}_{\mu\nu} = \frac{1}{2} \eta^{\alpha\beta} h_{\mu\nu,\alpha\beta} = \frac{1}{2} \square_{SR} h_{\mu\nu}. \quad (2.14)$$

Where we have defined the Special Relativity (SR) D'Alembertian differential operator as:  $\square_{SR} = \eta^{\mu\nu} \partial_\mu \partial_\nu$ . We can plug eq. 2.14 into

eq. 2.5, with  $\Lambda = 0$  and up to first order in  $h_{\mu\nu}$ , we get the linearized Einstein field equations for a system of harmonic coordinates:

$$\square_{SR} h_{\mu\nu} = -16\pi G \left( T_{\mu\nu} - \frac{\eta_{\mu\nu}}{2} T^\alpha_\alpha \right), \quad (2.15)$$

$$h^\alpha_{\mu,\alpha} = \frac{1}{2} h^\alpha_{\alpha,\mu}. \quad (2.16)$$

Recall raising and lowering indices is done with the Minkowski metric. The tensor  $S_{\mu\nu}$  encodes the behavior of the source of GWs. One could also plug eq. 2.14 into eq. 2.1, with  $\Lambda = 0$ , and express it in terms of the trace reversed perturbation:  $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h^\alpha_\alpha \eta_{\mu\nu}$ . Then, we obtain similar and simpler equations at the cost of a more complex perturbation:<sup>4</sup><sup>5</sup>

$$\square_{SR} \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu} \quad (2.17)$$

$$\bar{h}_{\mu\nu}{}^\nu = 0. \quad (2.18)$$

We can easily recover the conservation equation for the stress-energy tensor. We take the divergence of eq. 2.17 and use eq. 2.18 to get:

$$T_{\mu\nu}{}^\nu = 0. \quad (2.19)$$

Let us look at what sorts of solutions come out of these

## 2.2 Homogeneous solutions

The first step is to consider the homogeneous solution, as all solutions will involve these terms. Setting  $T_{\mu\nu} = 0$  or  $S_{\mu\nu} = 0$  yields the wave equation in the absence of sources,

$$\square_{SR} h_{\mu\nu} = 0. \quad (2.20)$$

The De-Donder gauge (eq. 2.16) and the remaining gauge freedom<sup>6</sup> restricts the possible forms of this solution to having only helicity  $\pm 2$  physically significant components (see Weinberg (1972)). Let us look at generic forms of the solution. The metric  $h_{\mu\nu}$  ought to be real-valued, thus we seek real solutions of the form

<sup>4</sup> Note that the inverse change of variables is just:  $h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2} \bar{h}^\alpha_\alpha \eta_{\mu\nu}$ .

<sup>5</sup> We eliminate the trace of the stress-energy tensor by using:  $R = 8\pi G T^\alpha_\alpha$ . We can write  $R^{(1)} = -\frac{1}{2} \square_{SR} \bar{h}^\alpha_\alpha$ , in de Donder gauge, which is precisely the extra term dropping out of  $\square_{SR} h_{\mu\nu}$  when we express it in terms of  $\bar{h}_{\mu\nu}$ .

<sup>6</sup> Associated to changes in coordinates of the form  $x^\mu \rightarrow x^\mu + \xi^\mu$  with  $\square_{SR} \xi^\mu = 0$ .

$$h_{\mu\nu} = \varepsilon_{\mu\nu} e^{ik \cdot x} + \varepsilon^*_{\mu\nu} e^{-ik \cdot x}, \quad (2.21)$$

where  $\varepsilon_{\mu\nu}$  is the polarization tensor,  $k$  is the wave vector and we define:  $k \cdot x \equiv \eta_{\mu\nu} k^\mu k^\nu = k_\mu x^\mu$ . The polarization tensor is a symmetric rank-2 tensor, since  $h_{\mu\nu}$  is. Substituting eq. 2.21 into  $\square_{SR} h_{\mu\nu} = 0$  gives  $k_\mu k^\mu \equiv k^2 = 0$ .<sup>7</sup><sup>8</sup> From eq. 2.16 we have

$$\varepsilon^\mu_{\nu} k_\mu = \frac{1}{2} \varepsilon^\alpha_{\alpha} k_\nu. \quad (2.22)$$

As said at the beginning of this subsection, we still have some remaining gauge freedom, which we now fix, choosing the following coordinate change  $x^\mu \rightarrow x^\mu + \zeta^\mu$  where:

$$\zeta^\mu = i A^\mu e^{ik \cdot x} = -i A^{*\mu} e^{-ik \cdot x} \quad (2.23)$$

In the new coordinates, the tensor perturbation takes the form:

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \zeta_\mu}{\partial x^\nu} - \frac{\partial \zeta_\nu}{\partial x^\mu} = \varepsilon'_{\mu\nu} e^{ik \cdot x} + \varepsilon'^*_{\mu\nu} e^{-ik \cdot x}. \quad (2.24)$$

with

$$\varepsilon'_{\mu\nu} = \varepsilon_{\mu\nu} + k_\mu A_\nu + k_\nu A_\mu \quad (2.25)$$

Equations -eq. 2.22 and -eq. 2.25 reduce the free components in the polarisation tensor to just two. Additionally, these equations conspire to yield a traceless polarisation tensor  $\varepsilon^\alpha_{\alpha} = 0$ , with  $\varepsilon_{0\mu} = 0$  (see Carroll (2019)). This in turn implies that the metric perturbation itself is traceless, and thus becomes equal to its trace-reversed counterpart. This is the so-called transverse traceless gauge:

$$h_{0\mu} = 0, \quad h^\alpha_{\alpha} = 0, \quad h_{\mu\nu}{}^\nu = 0. \quad (2.26)$$

The metric perturbation in this gauge is written as:  $h^{TT}_{ij}$ . In empty space we can write the two independent components of the metric perturbation as:

$$h^{TT}_{ab}(t, \mathbf{x}) = \begin{pmatrix} h_+ & h_x \\ h_x & -h_+ \end{pmatrix}_{ab} \cos[\omega t - \mathbf{k} \cdot \mathbf{x}], \quad (2.27)$$

where  $h_+$  and  $h_x$  are the so-called plus and cross polarisations.

<sup>7</sup> Assuming non-zero perturbation  $h_{\mu\nu}$  of course.

<sup>8</sup> This is saying that the wavevector for the wave is null, thus that the wave propagates at the speed of light.

## 2.3 Inhomogeneous solutions

With the homogeneous part of eq. 2.15 accounted for, we are now ready to derive the inhomogeneous part. The solution of eq. 2.15 in the presence of a source term will be heavily inspired by the analogous problem in electromagnetism, where one learns that any linear differential equation:

$$\hat{\mathcal{L}}\phi = J \quad (2.28)$$

is solved using green's functions. If we define the following retarded Green's function

$$\mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) = -2\pi\delta((x - x')^2)\Theta(t > t'), \quad (2.29)$$

which satisfies,

$$\square_{SR}\mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) = -2\pi\delta^4(x^\mu - x'^\mu). \quad (2.30)$$

Then, the solution to

$$\square_{SR}h_{\mu\nu} = -16\pi GS_{\mu\nu} \quad (2.31)$$

is given by

$$h_{\mu\nu}(x) = 8G \int d^4x' \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) S_{\mu\nu}(x'). \quad (2.32)$$

Indeed, plugging in eq. 2.32 into eq. 2.31,

$$\begin{aligned} \square_x h_{\mu\nu} &= 8G \int d^4x' \left( (\underbrace{\square_x \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu)}_{=-2\pi\delta(x^\mu - x'^\mu)}) S_{\mu\nu}(x') \right) = -16\pi GS_{\mu\nu}(x). \end{aligned} \quad (2.33)$$

We can perform the  $x'^0 = t'$  integration in eq. 2.32, with the Dirac delta function, which sets  $t'$  to be the retarded time,  $t' = t - |x^\mu - x'^\mu| = t_r$ , i.e:

$$h_{\mu\nu}(x) = 8G \int d^3y S_{\mu\nu}(t, y) dt' \frac{\delta(t' - (t - |\mathbf{x} - \mathbf{y}|))}{2|\mathbf{x} - \mathbf{y}|} \quad (2.34)$$

One gets the parallel solution to the trace reversed equation eq. 2.17 by swapping  $S$  with  $T$  and all  $h$  with  $\bar{h}$ .

$$h_{\mu\nu}(\mathbf{x}, t) = 4G \int d^3y \frac{S_{\mu\nu}(t - |\mathbf{x} - \mathbf{y}|, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|}. \quad (2.35)$$

We can interpret  $h_{\mu\nu}(\mathbf{x}, t)$ , as the perturbation obtained by summing up all the radiation from point-sources located at  $(\mathbf{x} - \mathbf{y}, t_r)$  on the past light cone. Put differently  $h_{\mu\nu}(\mathbf{x}, t)$  is the gravitational radiation produced by the source  $S_{\mu\nu}$ . Additionally, the form of the time argument of the source tensor, imposed by the definition of the Green's function, shows that the radiation propagates with velocity  $= 1 = c$ . The solution given by eq. 2.35 satisfies the harmonic coordinate condition of eq. 2.15, since

$$T_{\mu\nu;\mu} = 0 \Rightarrow T_{\mu\nu,\mu} + \underbrace{\Gamma\Gamma}_{\text{non-linear}} = 0 \Rightarrow T_{\mu\nu,\mu} = 0, \quad (2.36)$$

where we have ignored the non-linear terms encoded in the Levi-Civita connections. Then,

$$S^{\mu\nu}_{,\mu} = \partial_\mu \left( T_{\mu\nu} - \frac{\eta^{\mu\nu}}{2} T^\alpha_\alpha \right) = -\frac{\eta^{\mu\nu}}{2} T^\alpha_{\alpha,\mu}. \quad (2.37)$$

Also

$$S^\mu_\mu = T^\mu_\mu - \frac{\delta^\mu_\mu}{2} T^\alpha_\alpha = -T^\alpha_\alpha. \quad (2.38)$$

Thus

$$\begin{aligned} S^{\mu\nu}_{,\mu} &= \frac{1}{2} S^\alpha_{\alpha,\mu} \eta^{\mu\nu} \\ S^\mu_{\nu,\mu} &= \frac{1}{2} S^\alpha_{\alpha,\nu}. \end{aligned} \quad (2.39)$$

Then

$$\begin{aligned}
h^\mu_{\nu,\mu} &= 8G \frac{\partial}{\partial x^\mu} \int d^4x' \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) S^\mu_\nu(x') \\
&= 8G \int d^4x' \frac{\partial \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu)}{\partial x^\mu} S^\mu_\nu(x') \\
&= -8G \int d^4x' \frac{\partial \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu)}{\partial x'^\mu} S^\mu_\nu(x') \\
&= \underbrace{8G \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) S^\mu_\nu(x')}_{=0 \text{ at the boundary}} + 8G \int d^4x' \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) \frac{\partial S^\mu_\nu(x')}{\partial x'^\mu} \\
&= 8G \int d^4x' \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) \frac{1}{2} \frac{\partial S^\alpha_\alpha(x')}{\partial x'^\mu} \\
&= \dots \text{ repeat in reverse} \\
&= \frac{\partial}{\partial x^\mu} \left\{ 8G \int d^4x' \mathcal{G}_{\text{ret}}(x^\mu - x'^\mu) \frac{1}{2} S^\alpha_\alpha(x') \right\} \\
&= \frac{1}{2} h^\alpha_{\alpha,\mu}.
\end{aligned} \tag{2.40}$$

## 2.4 Gravitational Wave Sources

Now that we have the general form of the solutions to the linearized Einstein equations, we can proceed to the analysis of sources of GWs. The first step is to analyse the equations in the frequency domain. We will use the following notation:

$$\mathcal{F}_t[\phi](\omega, \mathbf{x}) = \int dt \phi(t, \mathbf{x}) e^{i\omega t} \tag{2.41}$$

$$\mathcal{F}_\omega^{-1}[\phi](t, \mathbf{x}) = \int \frac{d\omega}{2\pi} \mathcal{F}_t[\phi](\omega, \mathbf{x}) e^{-i\omega t} \tag{2.42}$$

Let us look at frequency domain of the solution given in eq. 2.35.

We have

$$\begin{aligned}
\mathcal{F}_t[h_{\mu\nu}](\omega, \mathbf{x}) &= \int dt h_{\mu\nu}(t, \mathbf{x}) e^{i\omega t} \\
&= 4G \int d^3y dt \frac{S_{\mu\nu}(t - |\mathbf{x} - \mathbf{y}|, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} e^{i\omega t} \\
&= 4G \int d^3y dt_r \frac{S_{\mu\nu}(t_r, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} e^{i\omega t_r} e^{i\omega |\mathbf{x} - \mathbf{y}|} \\
&= 4G \int d^3y \frac{\mathcal{F}_{t_r}[S_{\mu\nu}](\omega, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} e^{i\omega |\mathbf{x} - \mathbf{y}|}.
\end{aligned} \tag{2.43}$$

We can now apply various approximations to this form of the perturbation:

1. Consider a compact source, centered on the origin of our coordinate system. This implies that  $S_{\mu\nu}$  is peaked and compactly supported in a region around the origin. We look at the radiation only in the so called *wave zone*, at a distance  $r = |\mathbf{x}|$  much larger than the dimensions of the source  $R = |\mathbf{y}|_{\max}$ .
2. We assume that  $r \gg \frac{1}{\omega}$ , i.e. long wavelengths don't dominate.
3. We assume that  $r \gg \omega R^2$ , i.e. the ratio of  $R$  to the wavelength is not comparable to the ratio of  $r$  to  $R$ . Using this approximation we can write:

$$\begin{aligned}
|\mathbf{x} - \mathbf{y}| &= r \left( 1 - 2\hat{\mathbf{x}} \cdot \mathbf{y} + \frac{\mathbf{x}'^2}{r^2} \right)^{1/2}, \\
|\mathbf{x} - \mathbf{y}| &\approx r \left( 1 - \frac{\hat{\mathbf{x}} \cdot \mathbf{y}}{r} \right) \quad \text{with: } \hat{\mathbf{x}} = \frac{\mathbf{x}}{r}.
\end{aligned} \tag{2.44}$$

If we additionally further separate the scales in the following way

<sup>9</sup>

$$r \gg [\frac{1}{\omega}, \omega R^2] \gg R, \tag{2.45}$$

Then eq. 2.43 becomes much simpler<sup>10</sup>:

$$\mathcal{F}_t[h_{\mu\nu}](\omega, \mathbf{x}) = 4G \frac{e^{i\omega r}}{r} \int d^3y \mathcal{F}_t[S_{\mu\nu}](\omega, \mathbf{y}). \tag{2.46}$$

<sup>9</sup> We just add the condition that  $\frac{1}{\omega} \gg R$ , that is the source radius is much smaller than the wavelength.

<sup>10</sup> The approximations all conspire to be able to neglect the  $\mathbf{y}$  dependence of  $\frac{e^{i\omega|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}$  in the integral.

Now, let us look at the Fourier transform of the source term. By definition, we have

$$\mathcal{F}_t[S_{\mu\nu}](\omega, \mathbf{y}) = \mathcal{F}_t[T_{\mu\nu}](\omega, \mathbf{y}) + \frac{1}{2}\eta_{\mu\nu}\mathcal{F}_t[T^\alpha_\alpha](\omega, \mathbf{y}) \quad (2.47)$$

Thus, we can study the Fourier transform of  $T_{\mu\nu}$ ,  $\mathcal{F}_t[T_{\mu\nu}]$  instead. Using the conservation equation eq. 2.19 in Fourier  $t$ -space, we obtain:

$$-\mathcal{F}_t[T_{i\mu}]^i = i\omega\mathcal{F}_t[T_{0\mu}] \quad (2.48)$$

This equation becomes algebraic when we further Fourier transform in  $\mathbf{x}$ -space:

$$\hat{T}_{\mu\nu}(k^\alpha) = \hat{T}_{\mu\nu}(\omega, \mathbf{k}) = \int d^3\mathbf{y} \mathcal{F}_t[T_{\mu\nu}](\omega, \mathbf{y}) e^{i\mathbf{k}\cdot\mathbf{y}} \quad (2.49)$$

Then the conservation equation becomes:

$$k^\mu T_{\mu\nu}(\omega, \mathbf{k}) = 0 \quad (2.50)$$

These four equations relate the time components of  $\hat{T}_{\mu\nu}$  to the spatial ones. We now apply eq. 2.48 to itself to obtain:

$$\mathcal{F}_t[T_{ij}]^{ij} = -\omega^2\mathcal{F}_t[T_{00}] \quad (2.51)$$

which, when multiplied by  $x_m x_n$ , and integrated over  $\mathbf{x}$  gives<sup>11</sup>

$$\int d\mathbf{x} \mathcal{F}_t[T_{mn}](\omega, \mathbf{x}) = -\frac{\omega^2}{2} \int d\mathbf{x} x_m x_n \mathcal{F}_t[T_{00}](\omega, \mathbf{x}). \quad (2.54)$$

Notice that the integral on the RHS is in fact the Fourier transform of the quadrupole moment tensor of the energy density:

$$I_{mn} = \int x_m x_n T_{00}(t, \mathbf{x}) d^3x. \quad (2.55)$$

We call its Fourier transform  $\hat{I}_{mn}(\omega)$  and we can finally rewrite eq. 2.46 as:

$$\mathcal{F}_t[h_{\mu\nu}](\omega, \mathbf{x}) = -2G\omega^2 \frac{e^{i\omega r}}{r} (\hat{I}_{mn}(\omega) + \frac{1}{2}\eta_{\mu\nu}\hat{I}^n_n(\omega)) \quad (2.56)$$

<sup>11</sup> Two integrations by parts cancel the  $x_m x_n$  term in the LHS and since boundary terms are 0 (the source is finite) we have

$$\int dx x_m x_n \mathcal{F}_t[T_{ij}]^{ij} = \int dx x_m x_n^{ij} \mathcal{F}_t[T_{ij}] \quad (2.52)$$

The hessian of  $x_m x_n$  is  $(\delta_m^i + \delta_n^i)(\delta_n^j + \delta_m^j)$ , but since  $T_{\mu\nu}$  is symmetric the integral is

$$\int dy 2\mathcal{F}_t[T_{mn}](\omega, \mathbf{y}) \quad (2.53)$$

Going back  $t$ -space we have:

$$\begin{aligned}
h_{\mu\nu}(t, \mathbf{x}) &= -\frac{G}{\pi r} \int d\omega e^{-i\omega(t-r)} \omega^2 (\hat{I}_{mn}(\omega) + \frac{1}{2} \eta_{\mu\nu} \hat{I}_n(\omega)) \\
&= \frac{G}{\pi r} \frac{d^2}{dt^2} \int d\omega e^{-i\omega(t_r)} (\hat{I}_{mn}(\omega) + \frac{1}{2} \eta_{\mu\nu} \hat{I}_n(\omega)) \quad (2.57) \\
&= \frac{2G}{r} \frac{d^2}{dt^2} (I_{mn}(t_r) + \frac{1}{2} \eta_{\mu\nu} I_n(t_r))
\end{aligned}$$

This equation has a nice physical interpretation. The gravitational wave produced by a non-relativistic source is proportional to the second derivative of the quadrupole moment of the energy density at the time  $t_r$  where the past light cone of the observer intersects the source. The nature of gravitational radiation is in stark contrast to the leading electromagnetic contribution which is due the change in the *dipole* moment of the charge density. The change of the dipole moment can be attributed to the change in center of charge (for Electromagnetism (EM)), or mass (for GR). While a center of charge is free to move around, the center of mass (of an isolated source) is fixed by the conservation of momentum, so the dipole moment is zero.

The quadrupole moment, on the other hand, is sensitive to the shape of the source, which a gravitational system can modify. We observe that the quadrupole radiation is sub-leading when compared to dipole radiation. Thus, on top of the much smaller coupling constant, gravitational radiation is also weakened by this fact, and thus is usually orders of magnitude weaker than electromagnetic radiation.

In summary, any object that is modifying its shape is a source of GWs. All orbiting systems therefore are sources of GWs. However, as outlined above, only significant ‘changes in shape’ have a chance to be detectable. We will now explore these phenomena.

#### 2.4.1 Compact binaries

How could one construct a very powerful source of GWs? An example would be that of two highly-massive bodies (such that  $T_{00}$  is large) that orbit each other. Furthermore, in order for the GWs emitted by the system to be detectable they must orbit close enough to each other that their quadrupole moment is large. For

these very massive objects to be close enough for a small orbit, they have to be very compact. Assuming that is the case, a funny thing happens, as these objects orbit each other, they emit GWs , and in doing so they lose energy <sup>12</sup>. Thus, they slow down, and their orbit shrinks. This continues until the orbit is so small that the objects merge into a single object. Of course this is a significant change of shape, and thus we have a constructed (if not all on purpose) a very powerful source of GWs . Such objects are called compact binaries.

Let us consider two massive bodies orbiting each other. The mass of each body is  $m_1$  and  $m_2$  respectively. The distance to the observer is  $r$  and the separation between the object is given by  $R$ . The setup is shown in Figure 2.1

To Leading Order (LO) in  $G$  the motion of the binary is described by a Newtonian circular orbit. In barycentric coordinates, such a system is in fact equivalent to effective one body <sup>13</sup> of reduced mass  $\mu = \frac{m_1 m_2}{M}$  orbiting in a gravitational potential of mass  $M = m_1 + m_2$ , at a distance  $R$  equal to the distance between the two bodies forming the binary. The orbit is characterized by a balancing act between the gravitational attraction due to the potential and the centripetal force due to the motion of the reduced mass.<sup>14</sup>

$$\frac{GM\mu}{(R)^2} = \frac{\mu u^2}{R}, \quad (2.58)$$

where  $u$  is the norm of tangential velocity, and  $R$  is the distance between the two objects. eq. 2.58 gives us the the norm of the velocity

$$u = \sqrt{\frac{GM}{R}}. \quad (2.59)$$

Thus, the orbital period is simply, in the case of

$$T = \frac{2\pi R}{u} = 2\pi \sqrt{\frac{R^3}{GM}}. \quad (2.60)$$

This also gives us angular frequency:

$$\Omega = \frac{2\pi}{T} = \sqrt{\frac{GM}{R^3}}. \quad (2.61)$$

<sup>12</sup> Of course this happens in every orbiting system just on a timescale that is negligible. Only systems which are massive enough to produce large amounts of radiation actually lose enough energy for it to matter.

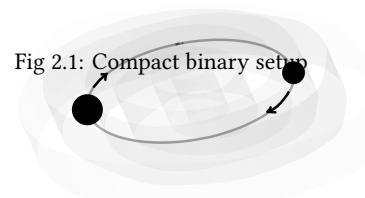


Fig 2.1: Compact binary setup

<sup>13</sup> This choice of wording is not a coincidence as we shall see

<sup>14</sup> Or equivalently between conserved angular momentum and the derivative of the potential energy.

Thus the EOM  $\mathbf{r}(t)$  of the reduced mass is described by a circular orbit in the  $x - y$  plane i.e.:

$$\mathbf{r}_{xy}(t) = \begin{pmatrix} R \cos(\Omega t) \\ R \sin(\Omega t) \end{pmatrix}. \quad (2.62)$$

We have not written the third  $z$  component of the trajectory as that can be set to zero, for a well-chosen set of inertial coordinates. The corresponding trajectory for both masses is simply given by applying the following scaling:

$$\mu\mathbf{r}(t) = m_1\mathbf{r}_1(t) = m_2\mathbf{r}_2(t). \quad (2.63)$$

We can straightforwardly obtain the 00 component of the energy-momentum tensor of the system <sup>15</sup> by using the EOM of the reduced mass (eq. 2.62):

$$T_{00}(t, \mathbf{x}) = \mu\delta(x_3)\delta(x_1 - R \cos(\Omega t))\delta(x_2 - R \sin(\Omega t)), \quad (2.64)$$

which when plugged into eq. 2.55, gives us the quadrupole moment of the rotating system:

$$\begin{aligned} I_{11} &= \mu R^2 \cos^2(\Omega t) = \mu R^2(1 + \cos(2\Omega t)) \\ I_{22} &= \mu R^2 \sin^2(\Omega t) = \mu R^2(1 - \cos(2\Omega t)) \\ I_{12} = I_{21} &= 2\mu R^2(\cos \Omega t)(\sin \Omega t) = \mu R^2 \sin(2\Omega t) \\ I_{33} &= 0. \end{aligned} \quad (2.65)$$

Thus, we have roughly:

$$I \sim \mu R^2 \cos(2\Omega t) \sim \mu R^2 \sim \mu \left( \frac{GM}{\Omega^2} \right)^{\frac{2}{3}} = \frac{Gm_1 m_2}{(m_1 + m_2)^{\frac{1}{3}}} \Omega^{-\frac{4}{3}}, \quad (2.66)$$

where we have expanded in small  $R$  in the second relation. The metric perturbation is then given by eq. 2.57 and has norm proportional to:

$$h \propto \frac{\ddot{I}}{r} \propto \frac{1}{r} \frac{Gm_1 m_2}{(m_1 + m_2)^{\frac{1}{3}}} \Omega^{\frac{2}{3}}. \quad (2.67)$$

<sup>15</sup> The energy density is the same for the effective one body system and the two body system.

## 2.4.2 BH Binaries

Taking the process described above to the limit of more massive and more compact objects one could imagine asking for the most dense objects possible to orbit each other. In GR this object is called a black hole (BH). It is a possible solution to the full fat EFE (eq. 2.1), in a static and isotropic universe, with point like mass at its center. Then the solutions to the equations eq. 2.1 have a unique form ([Birkhoff and Langer 1923](#)), called the Schwarzschild solution.

The solution is given by, in spherical coordinates:

$$d\tau^2 = \left[ 1 - \frac{R_S}{r} \right] dt^2 - \frac{dr^2}{1 - \frac{R_S}{r}} - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \quad (2.68)$$

where  $R_S$  is the Schwarzschild radius, given by:<sup>16</sup>

<sup>16</sup> we reinstate the speed of light, as we want to compute orders of magnitudes

$$R_S = \frac{2GM}{c^2}. \quad (2.69)$$

For stellar masses we have  $R_S = 2954$  m, quite a lot smaller than the radius of the sun  $\sim 7 \cdot 10^8$  m. Note that the metric given in eq. 2.68 exhibits singularities at both  $r = R_S$  and  $r = 0$ . The former is spurious, as it is an artefact of the coordinate choice, while the latter is real, in the sense that coordinate independent scalars formed from the metric (such as the ricci scalar) do in fact blow up as  $r \rightarrow 0$ . This encodes the non-spurious mass singularity that gives rise to the BH phenomena.

If we consider a comparable stellar mass BBH system, we can use the calculations of the previous section to compute the order of magnitude of the perturbation. eq. 2.61 in terms of the Schwarzschild radius gives us the orbital frequency

$$f = \frac{\Omega}{2\pi} = \frac{c}{2\pi} \sqrt{\frac{R_S}{R^3}}, \quad (2.70)$$

and the metric perturbation is given by eq. 2.67:

$$h = \frac{R_S^2}{R}. \quad (2.71)$$

Thus, if we consider BHs separated by ten times their Schwarzschild radius, observed at cosmological distances  $\sim 100$  Mpc, we obtain the values for the scales characterizing the binary system:

$$\begin{aligned} R_S &\approx 10^3 \text{ m}, \\ R &\approx 10^4 \text{ m}, \\ r &\approx 10^{23} \text{ m}. \end{aligned} \tag{2.72}$$

Such a system would have an orbital frequency of:

$$f \approx 10^2 \text{ Hz}, \tag{2.73}$$

and would perturb the metric on earth by:

$$h \approx 10^{-21}. \tag{2.74}$$

To have any hope of detecting such systems, one would need to have detectors sensitive to frequency spectra up to hundreds of Hertz, with a strain sensitivity of  $10^{-21}$ . This has been achieved at LIGO and other detectors, and has enabled the detection of such systems.

Non-spinning black holes orbiting around each other are the simplest binary system to model. However, they are extremely unlikely to occur in nature (realistic physical scenarios). The nature of BH formation implies that they necessarily also spin. Spinning black holes, also referred to as Kerr black holes, are characterized by a different metric tensor defining the spacetime surrounding them. The modelling of spinning black holes is still in its infancy and will not be treated it here.

There are many other binary systems that could feasibly be detected. Finite-size binaries such as Neutron Stars (NSs) result in tidal phenomena that are not present in the BH case. Also, non-binary systems, such as Cosmic strings, corresponding to early universe topological defects, could produce a detectable stochastic GW background.

## 3 Gravitational Wave Detection

Gravitational waves were first theorized by Einstein, accompanying his theory of general relativity. However, at the time it was not even clear whether they were an artifact of the theory or an observable phenomenon. Of course, as we have seen, GWs are predicted to be extremely weak, and thus there was initially little hope of ever detecting them. As evidence mounted that GR was indeed an incredibly accurate theory of gravity, the interest in GW detection grew. The first detectors, resonant antennas, failed in successfully detecting GWs, as they were simply not sensitive enough. The more modern laser based detectors were finally able to claim detection, by being extremely sensitive (the modern measurement precision at LIGO is equivalent to measuring the distance to Alpha Centauri 4.367 light-years, to the width of a human hair  $\sim 40 \mu\text{m}$ ). The first detection was made in 2015 ([B. P. Abbott et al. 2016](#)), marking the beginning of a new era in astronomy, and maybe even physics itself.

### 3.1 Laser interferometers

Laser-interferometer-based GW detectors are essentially scaled-up versions of the Michelson interferometer. The device, shown in [?@fig-fig-interfertrans](#), splits a beam of colimated coherent photons (laser) into two, letting them propagate along two axes in an L configuration, before being reflected back to the beam splitter. The two beams, when recombined, interfere with each other, constructively or destructively, depending on the nature of the medium they have traveled through and the length of each arm.

For a gravitational wave to produce a measurable effect on such an apparatus, all the optical equipment should be free to move and be isolated from external vibrations. This is achieved by hanging the mirror on precisely-tuned penduli. When the gravitational wave

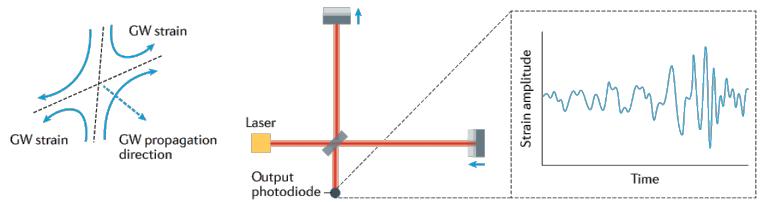


Fig 3.1: Principles of laser interferometer GW detectors

propagates through the apparatus, it induces a relative movement between the masses, changing the difference  $L_x - L_y$  in length of the detectors arms, by an amount  $\Delta L$ . This effect is exacerbated as the length of the arms  $L_x \sim L_y \sim L$  is increased. It follows that the strain is the normalized difference:

$$\frac{\Delta L}{L} = [(1/2)(1 + \cos^2(\theta)) \cos(2\phi)] h_+ + [\cos(\theta) \sin(2\phi)] h_x \equiv h, \quad (3.1)$$

where  $\theta$  and  $\phi$  are the polar and azimuthal angles of the direction of propagation of the GW . The strain is decomposed into components,  $h_+$  and  $h_x$ , which are the plus and cross polarisations of the wave. The strain sensitivity can be further increased by bouncing the light back and forth multiple times along the arms of the interferometer. In current implementations this is achieved using Fabry - Perot cavities spanning the whole length of the arms.

### 3.1.1 LIGO and ground based observatories

LIGO implements the measurement apparatus described in the past section, and a schematic is shown in Figure 3.2. It consists of two detectors, both in the USA, one in Livingston Louisiana and one in Hanford Washington, a separation of 3000 km. Each LIGO detector has two 4 km long arms made up of 1.2 m-wide vacuum tubes, each operating as Fabry-Perot cavities. The light from an effective 750 kW laser bounces back and forth inside these arms around 300 times. Thus, the effective arm length becomes 1200 km, drastically improving sensitivity.

While LIGO was the first to provide detection data, it does not live in a vacuum. In fact for the first published detection, the data analysis teams from LIGO joined forces with those for Virgo, the European gravitational wave detector. The Virgo detector is located in Cascina, Italy, and is similar to LIGO . It has 3 km long arms and a different suspension system. It joined data collection

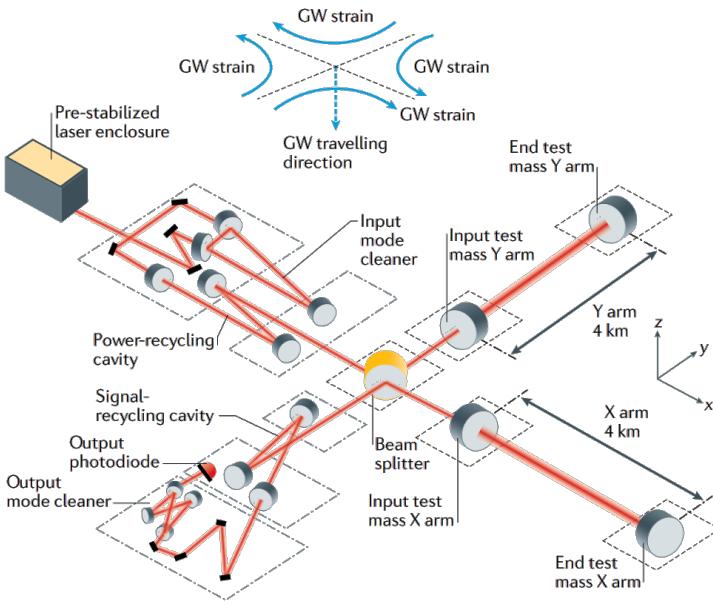


Fig 3.2: Schematic of LIGO detector

with LIGO in the second observing run in 2017. Finally, there are two other detectors that are now part of the collaboration, KAGRA in Japan and GEO600 in Germany. KAGRA is a similar detector to Virgo in scale, but is built underground further isolating it from seismic activity .GEO600 is a smaller detector, with 600 m long arms, built with the objective of being a test bed for future detectors.

Although they exhibit impressive sensitivity, ground based detectors are limited, since they are subject to seismic activity, which can cause vibrations that limits the lower bound of the frequency sensitivity of the detectors. The optics and other higher frequency noise imposes an upper bound to the frequency. In particular ground based detectors are only sensitive to frequencies in the following rough range:

$$[10^0, 10^4] \text{ Hz.} \quad (3.2)$$

In turn this implies that they are sensitive to the comparable, stellar mass <sup>17</sup> compact binaries, which are currently believed to be the most common type of binary systems in the universe. Of course this is not an accident, as these detectors were purpose built for providing evidence of GW in general. In fact, they have been quite successful in this regard, with upwards of 90 detections of GW to date.

<sup>17</sup> and a bit bigger  $\sim 1 - 10^3 M_{\odot}$

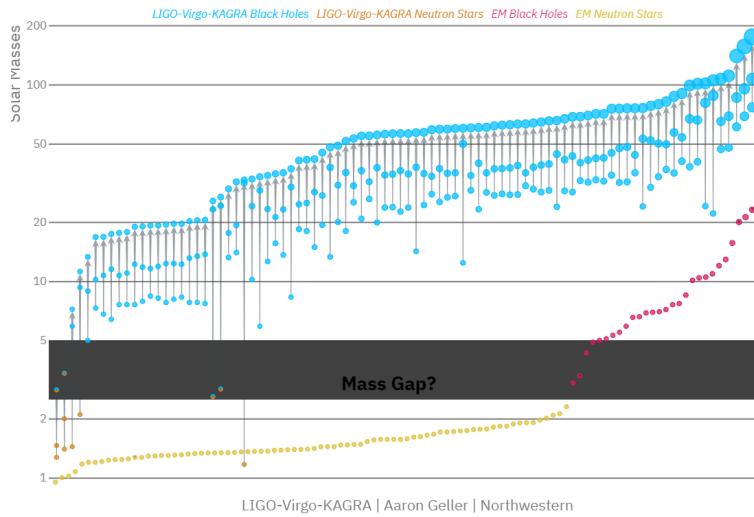


Fig 3.3: Mass distribution of Detected GW emitting binaries. Credit: Visualization: LIGO-Virgo-KAGRA / Aaron Geller / Northwestern

Figure 3.3 shows the current library of compact binary systems detected by LIGO -Virg-Kagra, along with those detected through non-GW means. The highlighted band denotes the assumed mass gap that should exist between the neutron star and black hole mass distributions. Note that most of the detections are compatible with the presence of such a gap. However, there are a few detections that are not, and these are particularly interesting. These are binaries with masses too light to be black holes, but too heavy to be neutron stars. A lot of research is ongoing to try and understand what these objects are, and if they are not just measurement artefacts.

### 3.1.2 LISA

There are many planned next generation GW detectors, from the Einstein telescope to Cosmic explorer. The most interesting one however is LISA, the first genuine space based GW observatory. It is a joint project between ESA and NASA, and is scheduled to launch in 2037. In Figure 3.4 we see it will consist of three satellites, in triangular formation, at a distance of 2.5 million kilometers of each other. Each of the three LISA spacecraft contains a test mass free-falling within its body. Each spacecraft then measures the distance to its neighboring test mass thus forming a space based laser interferometer.

There are a few key differences with ground based detectors. The first is the size of the arms means it will be sensitive to frequency

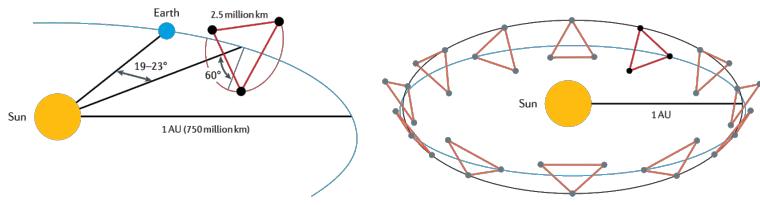


Fig 3.4: Schematic of LISA detector

ranges such as:

$$[10^{-4}, 10^0] \text{ Hz.} \quad (3.3)$$

This means that it will be sensitive to the most massive binaries in the universe, supermassive black holes, such as those found at the centers of galaxies.

The second is that it will be completely isolated from seismic activity, and thus has a much lower noise floor. This combined with the arm length means that LISA is also predicted to be able to observe Extreme Mass Ratio Inspirals (EMRIs), where stellar compact objects ( $< 60M_{\odot}$ ) orbit around a massive black hole of around  $10^5 M_{\odot}$  or more. In this configuration the small mass is in a slowly decaying orbit, and thus can spend around  $10^5$  orbits and between a few months and a few years in the LISA sensitivity band before merging. This will provide us with a very interesting new source of GW , that also allows us to test new approximations, such as those based on mass ratio expansions: Gravitational Self-Force (GSF) .

### 3.2 Matched filtering

The signals measured by the laser interferometers such as LIGO , provide a measurement of strain over time. This waveform essentially measures the deformation of spacetime when a GW passes through a detector. While some GW signals observed in laser interferometers are clearly identified as such, mainly due to their extraordinary power output, weaker signals are harder to identify. In fact, most of the detections have happened under the noise floor of the detectors. How is this possible? Matched filtering and waveform generation.

Consider some detector output signal  $s$  consisting of noise  $n$  and a possible true GW signal  $h_i$  (or signals, making up a family  $\{h_i\}$ ).

The nature of the noise is unpredictable and necessarily probabilistic in nature. We can, however, model the noise as a random variable  $n$  that follows a probability distribution  $P_n$  dependent on variables that we consider and have measured to be significantly impactful.

In such a probabilistic setup one could try to find the probability that the signal  $h_i \in \{h_i\}$  has been detected, conditional on the actual data  $s$ . This however is impractical as it requires *a priori* knowledge on the family of possible signals, information that is inaccessible. Instead, we devise a functional  $\mathcal{F}[s, h_i]$ , (a statistic) dependent on both the data and the supposedly present GW signal. We can then choose a threshold  $\rho^*$  above which we claim detection of the signal  $h_i$ , i.e. if  $\mathcal{F}[s, h_i] > \rho^*$ .

The functional that is used is defined in terms of the symmetric inner product  $\langle s, h_i \rangle$ , that correlates the data  $s$  with the template  $h_i$ , weighting by detector frequency sensitivity  $S_n(f)$ :

$$\langle g, h \rangle = 2 \int_{-\infty}^{\infty} \frac{\hat{g}^*(f)\hat{h}(f)}{S_n(|f|)} df = 4\Re \int_0^{\infty} \frac{\hat{g}^*(f)\hat{h}(f)}{S_n(f)} df, \quad (3.4)$$

where  $\hat{g}^*$  is the conjugate of the frequency Fourier transform of  $g$ .  $S_n(f)$  is the *noise power spectral density*, characterising the detector sensitivity defined by:

$$\overline{\hat{n}^*(f_1)\hat{n}(f_2)} = \frac{1}{2}\delta(f_1 - f_2)S_n(f_1)\Theta(f_1), \quad (3.5)$$

where the bar denotes the average over the noise realisation,  $\delta(f_1 - f_2)$  is the Dirac delta function, and  $\Theta(f_1)$  is the Heaviside step function. We can now define the *signal to noise ratio* for the signal when filtered by  $h_i$  as<sup>18</sup>

$$\rho(s, h_i) = \frac{\langle s, h_i \rangle}{\text{rms}\langle n, h_i \rangle} = \frac{\langle s, h_i \rangle}{\sqrt{\langle h_i, h_i \rangle}}. \quad (3.6)$$

where rms denotes the root-mean-square. This is the functional  $\mathcal{F}[s, h_i]$  we are after and we claim detection if  $\rho(h_i) > \rho^*$ . The functional  $\mathcal{F}[s, h_i] = \rho(s, h_i)$ , provided that the noise is normally distributed, is provably the optimal filtering technique, in the sense that it minimizes the probability of false positives. For example, if we have a fully Gaussian noisy signal  $s = n$  then  $\rho$  is a normal

<sup>18</sup> The equality is due to the fact that we average over all realisations of noise, which averages to uniform:  $\overline{\langle h_i, n \rangle \langle n, h_i \rangle} = \langle h_i, h_i \rangle$ .

variable with zero mean and unit variance. If a signal  $h_i$  is present, over some again Gaussian noise  $n$ , then it can be proven that  $\rho$  is still normal but centered on  $\sqrt{\langle h_i, h_i \rangle}$  with unit variance.

The threshold can be set to correspond to a desired false positive rate  $\mathcal{F}$ :<sup>19</sup>

$$\mathcal{F} = \sqrt{\frac{1}{2\pi}} \int_{\rho^*}^{\infty} e^{-\frac{1}{2}\rho^2} d\rho = \frac{1}{2} \operatorname{erfc}\left(\frac{\rho^*}{\sqrt{2}}\right), \quad (3.7)$$

or correct detection rate  $\mathcal{D}$ :

$$\mathcal{D} = \frac{1}{2} \operatorname{erfc}\left[\frac{\rho^* - \sqrt{\langle h_i, h_i \rangle}}{\sqrt{2}}\right]. \quad (3.8)$$

Given a signal and a template, we have described a procedure to determine if the signal is present or not. The physical output of a detection is therefore bounded from above by the physics content of the template. The matched filtering approach is contingent on having a waveform that corresponds to the physical process that is emitting the signal. Usually, these processes can be parametrized by a set of parameter. In the case of a compact binary these parameters are for example, the mass ratio of the binary, the orbit eccentricity, the possible spin etc. The filter waveform must then naturally depend on these parameters, be it heuristically or physically (i.e., from first principles).

### 3.3 PTA

A PTA consists of an array of several pulsars whose pulse periods (or Time of Arrival (ToA)) are monitored with very high precision by one or more radio telescopes. Pulsars, from **pulsating** radio source, are though to be highly magnetized rotating neutron stars. These neutron stars emit strong electromagnetic radiation from their magnetic poles, which when not aligned with rotational poles, emit in the manner of a lighthouse. The measured pulses are extremely regular in nature, with periods ranging from a few milliseconds to a few seconds. Over a year, the variation in the pulse period is on the order of a few microseconds ([Detweiler 1979](#)). We can therefor use pulsars as clocks, and infer their variation

<sup>19</sup> The complementary error function is given by  $\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-t^2} dt = 1 - \operatorname{erf}(z)$ .

from their normal cycle time. This methodology is shown in Figure 3.5.

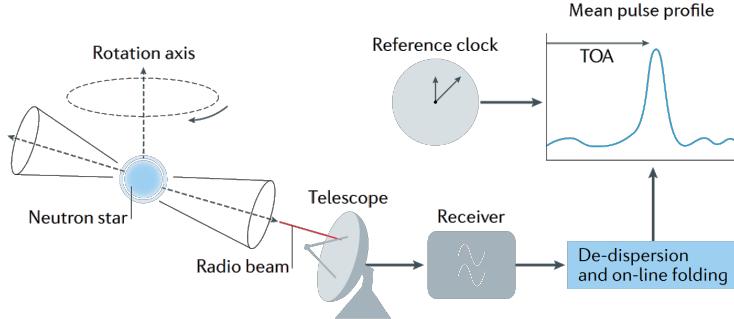


Fig 3.5: How pulsar timing arrays work

GWs propagating between the pulsar and the observer act as anisotropic media. This will cause the pulse to be delayed or sped up, depending on the relative direction of the wave, its polarisation axes and the direction of the pulsar. Given a collection of such pulsars with precisely known periods, evenly distributed in the celestial sphere, then one can use the relative delay of the pulses to measure the direction and nature of the wave. This setup is essentially a blown up version of LIGO with light-year scale arms instead of kilometer scale arms.

Currently, the set of pulsars suitable for this type of measurement is around 40 ([Maiorano, De Paolis, and Nucita 2021](#)). Many more pulsars have been measured but not all are suitably regular to be used in PTAs . The frequency range of sensitivity of PTAs has a lower bound set by observation duration and higher bound set by cadence of measurement. Currently, the IPTA Hobbs et al. ([2010](#)) has been measuring for 13 years, with measurements being taken every week yielding a sensitivity in a drastically different band than that of LIGO and LISA :

$$[10^{-9}, 10^{-7}] \text{Hz} \quad (3.9)$$

The nature of sources that can be detected by PTAs is therefore different from that of LIGO and LISA. The PTA sensitivity is more suited to detect stochastic gravitational waves, such as those emitted by early-universe phenomena, or by supermassive black holes ( $10^6 - 10^9 M_\odot$ ) in the centers of galaxies <sup>20</sup>.

<sup>20</sup> more massive objects have longer orbital periods

## 4 Waveform Generation

The matched filtering approach introduced in the previous section motivates the development of tools for generating waveforms. The matched filtering technique is especially sensitive to the phase of the GW , necessitating precise determination of the waveform template's phase. These templates can be varied in physical content, but also in practicality. In all cases they need to be able to cover the set of parameters (the parameter space) that one is interested in exploring. Clearly, if the waveform has been generated purely heuristically, then the physical content of the detection is close to zero. From this approach, one can only deduce that a phenomenon producing a similar waveform was detected, but not much more. However, if the input parameters are physically meaningful, e.g. the mass ratio of a binary compact inspiraling system, then the waveform constructed purely from this input, if matched to the signal, tells us that a compact binary inspiraling system with this mass ratio has been detected. This is called parameter estimation, and is key to maximizing the physics output of a detection.

Crucially, the waveform generation must be able to cover the whole parameter space (for example all feasible mass ratios) to be useful, as one does *not* know in advance the parameters describing the system to be identified. Ideally waveform generation should be done in a way that is computationally efficient, as the number of parameters as well as their range to explore is large. In the end, waveform generation is currently performed in a hybrid way: some waveforms for specific parameter space points, are generated from first principles, while the rest of the parameter space waveforms are interpolated.

In this chapter we will describe the first-principles waveform generation tools currently being developed. More specifically we will explore techniques for compact binary waveforms. In this context waveforms may describe inspiralling or scattering binaries. Currently, only the inspiralling binaries have been detected, as these are the systems whose signal LIGO and other detectors

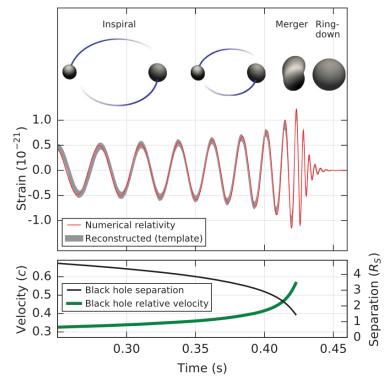
are most sensitive to (see Section 3.1.1), other than being by far the most common emitters of high-intensity GWs. As discussed in Section 2.4.1, orbiting and GW emitting compact objects will necessarily inspiral and merge at some point.

For such inspiraling binaries, the waveform features 3 distinct phases shown in : a first *inspiral phase*, a *merger phase* where a remnant compact body is produced as a result of the coalescence of the two objects, and a postmerger, or *ringdown phase*, where the remnant still emits gravitational radiation while settling to its new stable configuration. Each phase is characterized by different dynamical properties of the orbiting bodies and thus different frameworks are aimed at describing the corresponding features of the waveforms. Even the description of one phase involves different regimes, and thus more specific framework, additionally dependent on the type/initial state of the two orbiting bodies.

There is also hope to detect scattering binaries as argued in Mukherjee, Mitra, and Chatterjee (2021), but they are currently mainly used as a testbed for new theories making use of particle physics knowledge. Additionally, one might be able to extend their validity to inspiraling binaries (see Kalin and Porto 2020a).

Most of these techniques have two important parts. The first consists in the precise modelling of the evolution of the emitting system. In most cases, this boils down to obtaining the EOMs for the interacting bodies. For scattering, this can just be the amplitude of the event. The second part is extracting the variation of the metric due to the source movement at the measurement point, far away from the source, i.e., the waveform. Of course, the metric at the source is tightly coupled to the dynamics of the system, and thus the approximations conducted at the source also effects the way one extracts the waveform.

The only full, first-principles waveform techniques are NR and EOB<sup>21</sup>. NR is quite unique when compared to other methods. Let us look at it before proceeding to EOB and all the inspiral waveform techniques.



<sup>21</sup> There are a whole family of phenomenological frameworks that we will not discuss

## 4.1 Numerical Relativity

The most precise, first-principles waveform-generation technique to date is numerical relativity (Numerical Relativity (NR)). It was

developed as early as the 1960s ([Hahn and Lindquist 1964](#)). Initially and until relatively recently, these numerical methods were unusable because of numerical instabilities. The idea, however, is simple: the EFE are a system of coupled partial differential equations (PDEs) (albeit nonlinear) that can be solved numerically by discretizing the metric. The causes of instability is three-fold. First, the aforementioned non-linearities make the discretization particularly sensitive to extreme variations, such as those close to a BH, which means that one needs a fine mesh to control error buildup. Second, the gauge freedom establishes limits on the type of discretization procedures, so that spurious modes are not introduced. Third, for BHs one needs to consider what to do about the physical metric singularity, as infinity is not a number one can plug into a computer! On top of stability problems, when compact binary objects are considered, the scales involved further complicate things, as one needs enough resolution to resolve the objects, and the deformations on the metric, as well as a large enough volume to have reasonable orbits. This requires a huge amount of lattice nodes, meaning such simulations (even if stable) are extremely computationally expensive.

In 2005 ([Pretorius Pretorius 2005](#)) and a little later ([Campanelli et al. Campanelli et al. 2006](#)) made a breakthrough advance and were able to fully compute an inspiral-merger-ringdown sequence for a BBH . Pretorius made use of two techniques: Adaptive Mesh Refinement (AMR) and the Generalized Harmonic (GH) formalism. AMR , like its name suggests, is a technique based on a dynamical discretization procedure . The GH formalism, tackles the singularity problem validating the excision of the black hole, and crucially letting that region move. The combination of the procedures, were able to finally slay the proverbial numerical dragon.

These waveforms, such as the one in Figure 4.2, confirmed the predictions made in ([Buonanno and Damour Alessandra Buonanno and Damour 2000](#)) that the merger phase of a BBH would not be as complicated as previously thought. As numerical techniques evolved further, they were able to elucidate more and more complex behaviors.

Nowadays, it is the gold standard of waveform generation. Of course, it remains incredibly computationally expensive <sup>22</sup> and making any sizable bank of purely NR waveforms is seen as a fools endeavor. To date, only a couple hundred parameter space points have been simulated. Consequently, NR is used to complement, calibrate and extend the other waveform generation techniques.

<sup>22</sup> A single waveform with 10 orbits can take a week and tens of thousands of CPU hours

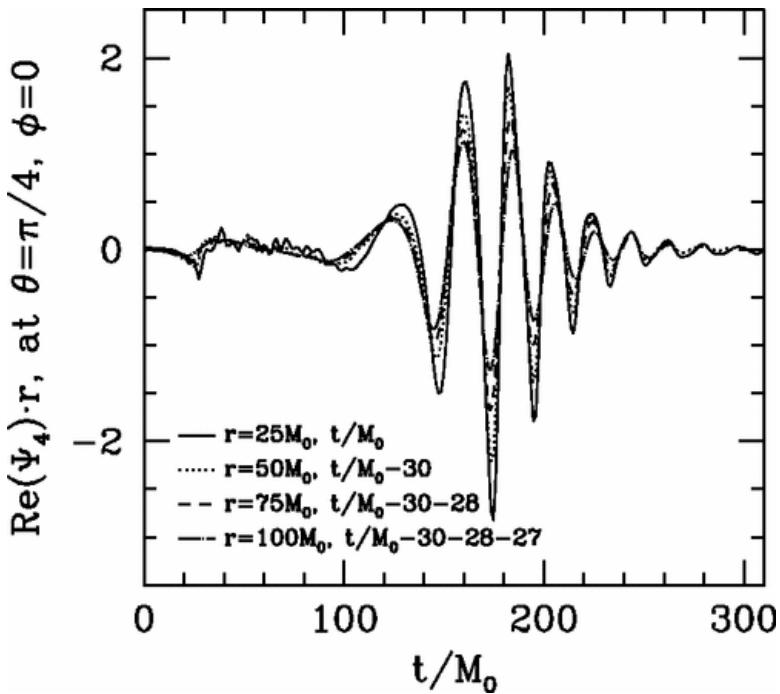


Fig 4.2: The first NR waveform, as presented in Pretorius (2005)

For example to detect the first signals LIGO and other detectors interpolated between NR-generated parameter space points to reach the ones needed for detection. However, with the advent of LISA on the horizon, a further problem will arise. NR results by nature only take into account the last few orbit cycles before merger. Any more than that and computational time becomes prohibitive. LISA, with its low noise floor, should be able to detect inspiraling but not yet merging binaries. Doing so will require waveform templates with hundreds of orbits, well out of reach of any modern NR package. Nonetheless, this technique has shown and will continue to show its worth, as no approximations other than discretization are made. The physical content of such waveforms is thus high.

## 4.2 Inspiral methods

Before continuing to the EOB method, let us look at the techniques that focus on the inspiral part of the waveform. There are broadly speaking three methods that fall in this category. The first we have already encountered at 1st order: PN . It is a perturbative expansion in both relative velocity (slow moving limit) and the coupling constant  $G$  (weak field limit) of the EFE . The second is the PM method, which is a perturbative expansion in *only* the coupling

constant  $G$ . At face value the PM method contains the PN one as a special case, however as at zeroth order, the PM approximation captures straight line motion, and is generally more expressive of scattering. The final approximation can be applied when the mass ratio of the object in question is large. Here one can expand around a fixed but *non-flat* metric, essentially considering one of the objects as a test-particle evolving on a background caused by the heavier object. This is called the GSF formalism.

### 4.2.1 Post-Newtonian

The Post-Newtonian (PN) expansion, is a perturbative expansion of EFE in the orbital regime, which at LO reproduces Newtonian gravity. Consider a system of orbiting bodies with typical values  $\bar{M}, \bar{R}, \bar{v}$  for masses, separations, and velocities. Then, the viral theorem states:

$$\langle T \rangle_\tau = -\frac{1}{2} \langle U \rangle \implies \bar{M} \bar{v}^2 \sim \frac{G \bar{M}^2}{\bar{R}} \quad (4.1)$$

thus we can consider the following equivalent small parameters:

$$\bar{v}^2 \sim \frac{G \bar{M}}{\bar{R}} \quad (4.2)$$

We therefore expand in both velocities and coupling constant. This can be achieved either by a classical (albeit quite complicated) perturbative treatment (see [Luc Blanchet 2014](#); [Luc Blanchet, Damour, and Iyer 1995](#)), or a more contemporary EFT-based approach coyly called NRGR ([Goldberger and Rothstein 2006](#); [Sturani 2021](#)), itself inspired by Non-Relativistic Quantum Chromodynamics (NRQCD). We call  $n$ PM the expansion carried out to  $\mathcal{O}(v^{2n})$ .

Regardless of method, there are a few shortcomings with this expansion. While it is well suited to describe the *motion* of objects in an orbital regime, it is not equally effective at describing gravitational radiation far from the source, as it exhibits divergences for  $r \rightarrow \infty$ <sup>23</sup>. This means that in the GW context, the PN expansion is in fact usually coupled to a weak-field approximation (see below), outside the near zone. This is the region surrounding the source, that is of small extent compared to the wavelength of the emitted radiation:

<sup>23</sup> In HEP speak this is called an infrared divergence

$$r \ll \lambda \quad (4.3)$$

The PN expansion can also be viewed as an expansion in small ‘retardations’. To see this, consider a slowly moving source, such that the time derivatives of the metric generated by this source will be smaller than the spatial derivatives by a factor  $\mathcal{O}(v)$ ,

$$\frac{\partial}{\partial t} = \mathcal{O}(v) \frac{\partial}{\partial x^i}, \quad (4.4)$$

or  $\partial_0 = \mathcal{O}(\epsilon)\partial_i$ . Thus, the d’Alembertian operator, applied to the metric is, to lowest order, a Laplacian operator

$$-\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \nabla^2 = [1 + O(\epsilon^2)] \nabla^2. \quad (4.5)$$

This means that retardation effects, due to the time derivative, are small corrections. In the PN expansion we are consequently trying to compute some quantity  $F(t - r)$ , such as a component of the metric, which is intrinsically a function of the retarded time  $t_r = t - r$ , from its expansion for small retardation  $r$ ,

$$F(t_r) = F(t) - r\dot{F}(t) + \frac{r^2}{2}\ddot{F}(t) + \dots \quad (4.6)$$

Importantly, the derivatives should bring down factors proportional to frequency  $\Omega$  of the emitted radiation. Thus, this is in fact an expansion in  $r\Omega \sim r/\lambda$ , which is only valid in the near zone as defined by eq. 4.3, and breaks down outside this zone, in the so-called *radiation zone*. A serious consequence of this fact is that the pure, naive PN expansion is unable to account for outgoing or incoming radiation, and thus unable to describe isolated systems. This is where the PM expansion comes in.

### 4.2.2 Post-Minkowski

The Post-Minkowski (PM) expansion is a weak field expansion of EFE, in the coupling constant  $G$ . It thus covers all orders in velocity of the PN expansion, at any fixed  $G$  order, see Figure 4.3. We usually call  $nPM$  the corresponding  $\mathcal{O}(G^n)$ -truncated expansion. Historically, it was derived with the objective of being the radiation zone partner to the PN expansion, as it provides a description of the propagation of the waves up to the observer. When used with

	0PN	1PN	2PN	3PN	4PN	5PN	6PN	7PN
1PM	$(1 + v^2 + v^4 + v^6 + v^8 + \dots)$							
2PM		$(1 + v^2 + v^4 + v^6 + v^8 + v^{10} + v^{12} + \dots)$						
3PM			$(1 + v^2 + v^4 + v^6 + v^8 + v^{10} + \dots)$					
4PM				$(1 + v^2 + v^4 + v^6 + v^8 + \dots)$				
5PM					$(1 + v^2 + v^4 + v^6 + \dots)$			
6PM						$(1 + v^2 + v^4 + \dots)$	$G^6$	
							$\vdots$	

the PN expansion in such a way, the PM expansion is used to solve the vacuum EFE. It is usually set up to compute a version of the metric perturbation that is exact:<sup>24</sup>

$$h_{\mu\nu} = \sqrt{-g}g^{\mu\nu} - \eta^{\mu\nu}. \quad (4.7)$$

The PM expansion is then given as

$$\sqrt{-g}g^{\mu\nu} = \eta^{\mu\nu} + G h_1^{\mu\nu} + G^2 h_2^{\mu\nu} + \dots, \quad (4.8)$$

meaning that:

$$h^{\mu\nu} = \sum_{n=1}^{\infty} G^n h_n^{\mu\nu}. \quad (4.9)$$

The EFE for this metric perturbation, when imposing De-Donder gauge conditions (eq. 2.16), is given by

$$\square_{SR} h^{\mu\nu} = -16\pi G h^{\mu\nu} \tau^{\mu\nu}. \quad (4.10)$$

where  $\tau^{\mu\nu}$  is the effective energy-momentum tensor given by:

$$\tau^{\mu\nu} = (-g)T^{\mu\nu} - \frac{1}{16\pi G}\Lambda^{\mu\nu}. \quad (4.11)$$

$\Lambda^{\mu\nu}$  contains all non-linear terms of the Einstein tensor, and its definition can be obtained by comparing with eq. 2.1. Thus, it depends on the metric and thus on  $h_{\mu\nu}$  in a highly non-linearly fashion. We can write it as:

$$\Lambda^{\mu\nu} = L^{\mu\nu}[h, h] + M^{\mu\nu}[h, h, h] + N^{\mu\nu}[h, h, h, h] + \mathcal{O}(h^5). \quad (4.12)$$

Fig 4.3: Comparison of PM and PM expansions, with currently attained orders encircled, and highlighted in red is the result of Z. Bern et al. (2019). Source Z. Bern et al. (2019)

<sup>24</sup> We have used  $h_{\mu\nu}$  to denote the linear perturbation, whereas  $h_{\mu\nu}$  is the all order metric perturbation.

Here the number of  $h$ 's in the argument of  $L, M$  and  $N$  indicates the order of the non-linear term. In a vaccuum eq. 4.10 reduces to

$$\square_{SR} h^{\mu\nu} = \Lambda^{\mu\nu}, \quad (4.13)$$

which we can solve iteratively order by order:

$$\begin{aligned} \square_{SR} h_1^{\mu\nu} &= 0, \\ \square_{SR} h_2^{\mu\nu} &= L^{\mu\nu}[h_1, h_1], \\ \square_{SR} h_3^{\mu\nu} &= M^{\mu\nu}[h_1, h_1, h_1] + L^{\mu\nu}[h_1, h_2] + +L^{\mu\nu}[h_2, h_1] \end{aligned} \quad (4.14)$$

This method can be supplemented by a multipolar expansion of the  $h_n^{\mu\nu}$ , truncated to the specific PM order we are interested in. This combined expansion was pioneered by (Bonnor and Rosenhead [Bonnor and Rosenhead 1959](#)), then formalised by (Thorne [Thorne 1980](#)), and finally applied in this present context by ([Blanchet L. Blanchet and Damour 1986](#)).

In terms of source dynamics, it is more appropriate (and uniquely able) to describe inherently relativistic phenomena, such as gravitational scattering. Its wider scope of applicability comes at the cost of a higher computational complexity, as at higher orders one encounters intricate families of relativistic integrals. Note that we have in fact computed the 1 PM approximation already in the Chapter 3 when we linearized the EFE. Higher orders are far more difficult to obtain, especially for sources such as BBH.

With the rise of amplitude techniques from high energy physics, both EFT and full theory results have been successful in extracting previously unreachable orders. Such techniques are perfectly suited to compute relativistic scattering as they were developed for this specific task. We will see this in the next chapter.

### 4.3 Effective One-Body

The second full-waveform generation technique is the Effective One-Body (EOB) formalism. It was initially developed by ([Buonanno and Damour A. Buonanno and Damour 1999](#)), as a non-perturbative re-summing of the PN expansion. It was later extended to include the merger and ringdown effects in ([Buonanno](#)

and Damour (Alessandra Buonanno and Damour 2000) and one can see the produced waveform in @. Thus, it became the first full waveform generating technique to actually produce a waveform. It has gotten progressively more powerful, as the PN expansions have gotten better. More recently PM results have also been included Damour (2016). Additionally, EOB can be combined with NR to produce a very accurate waveform, with less computational overhead than NR alone.

As its name suggests, the EOB framework combines all the PN, PM and NR data into a Hamiltonian describing the motion of a single body. It is inspired by solutions of two-body problems in Newtonian gravity which make use of barycentric coordinates to turn it into a one body problem. Further inspiration came from the two body problem in Quantum Electrodynamics (QED) (see Brezin, Itzykson, and Zinn-Justin 1970) .

Consider two bodies, with masses  $m_1$  and  $m_2$ , positions  $\mathbf{q}_1, \mathbf{q}_1$  and momenta  $\mathbf{p}_1$  and  $\mathbf{p}_2$  respectively. The system can be described by the following Hamiltonian:

$$\mathcal{H}_{\text{newton}}(\mathbf{q}_1, \mathbf{q}_1, \mathbf{p}_1, \mathbf{p}_2) = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} - \frac{Gm_1m_2}{|\mathbf{q}_1 - \mathbf{q}_1|} \quad (4.15)$$

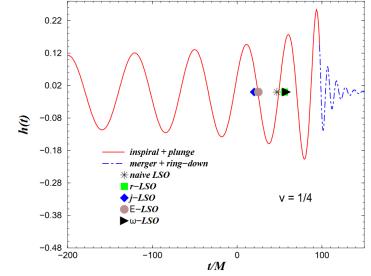
One then transforms this Hamiltonian into one describing the relative motion  $\mathbf{q} = \mathbf{q}_1 - \mathbf{q}_1$ . One can formalize this by defining a new set of generalized coordinates, through a canonical transformation. We can go even further and employ a generating function for the transformation, ensuring that Hamilton's equations remain unchanged in form (see Goldstein, Poole, and Safko (2002)). To describe relative motion, we use Jacobi coordinates ( $\mathbf{Q}, \mathbf{q}$ ):

$$\begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{pmatrix} \rightarrow \underbrace{\begin{pmatrix} \frac{m_1}{M} & \frac{m_2}{M} \\ 1 & -1 \end{pmatrix}}_A \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{q} \end{pmatrix} \quad (4.16)$$

If we now define the following generating function for the transformation, depending on the old positions coordinates  $\mathbf{q}_i$  and the new momenta  $\mathbf{P}_i$ :

$$f_{II}(\mathbf{q}_i, \mathbf{P}_i) = \mathbf{P}_i A^{ij} \mathbf{q}_j, \quad (4.17)$$

then the old momenta are given by



Note the use of Einstein summation

$$\mathbf{p}_i = \frac{\partial f_{II}}{\partial \mathbf{q}_i} = \mathbf{P}_i A^{ij} \iff \mathbf{P}_i = [A^{-1}]^{ij} \mathbf{p}_j. \quad (4.18)$$

Where we can directly invert this equation and obtain the new momenta as:

$$\begin{pmatrix} \mathbf{P} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{M} \\ \frac{m_2}{M} & -\frac{m_1}{M} \end{pmatrix} \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \end{pmatrix} \quad (4.19)$$

As expected, we obtain that one of the momenta describes the center of mass momentum  $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$ . The other positions are by construction given by  $\frac{\partial f_{II}}{\partial \mathbf{q}_1} = \mathbf{Q}$  and  $\frac{\partial f_{II}}{\partial \mathbf{q}_2} = \mathbf{q}$  i.e the same as equation eq. 4.16. Finally the new Hamiltonian  $\mathcal{H}_{\text{rel}}(\mathbf{Q}, \mathbf{q}, \mathbf{P}, \mathbf{p})$  can be directly obtained from the generating function:<sup>25</sup>

$$\mathcal{H}_{\text{rel}}(\mathbf{Q}, \mathbf{q}, \mathbf{P}, \mathbf{p}) = \mathcal{H}_{\text{newton}}(\mathbf{Q}, \mathbf{q}, \mathbf{P}, \mathbf{p}) + \cancel{\frac{\partial f_{II}}{\partial t}}^0 = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} - \frac{Gm_1m_2}{|\mathbf{q}|}, \quad (4.22)$$

where  $M = m_1 + m_2$  is the total mass, and  $\mu = \frac{m_1m_2}{M}$  is the reduced mass. If we change coordinates to a frame where the  $\mathbf{Q}$  is at rest, i.e.  $\mathbf{P} = 0$ , the Hamiltonian collapses to a one body Hamiltonian. The two-body problem we started with can now be thought of as the one body problem of a test-particle of mass  $\mu$  orbiting an external mass  $M$ .

In Newtonian gravity, a small change of coordinates makes the two body problem exactly solvable. In much the same way, EOB generalizes this framework to general relativity. The defining trick used above is to transform to ‘better’ coordinates. A first step is to demand that for straight line motion the Center of Momentum (COM) energy  $E_{\text{COM}} = P^0$ , where  $P$  is the 4-momentum of the entire-system, is equal to the Hamiltonian. We then have:<sup>26</sup>

$$\mathcal{H} = \sqrt{M^2(1 + 2\nu(\gamma - 1))}, \quad (4.24)$$

where  $\nu = \frac{\mu}{M}$  is the symmetric mass ratio, and  $\gamma = \frac{1}{\sqrt{1-\nu^2}} = \frac{p_1^\mu p_2^\mu}{m_1 m_2}$  is the relative Lorentz-factor. Note that for non-interacting systems we have  $p_\mu p^\mu = \mu^2$  and we can define the effective Hamiltonian:

<sup>25</sup> We just plug in the definition of the old momenta in terms of the new (eq. 4.18):

$$\mathbf{p}_1 = \frac{m_1}{M} \mathbf{P} + \mathbf{p}, \quad (4.20)$$

$$\mathbf{p}_2 = \frac{m_2}{M} \mathbf{P} - \mathbf{p}. \quad (4.21)$$

<sup>26</sup> In the COM frame, the spatial part of the total momentum is zero (by definition), and thus we have:

$$P_0^2 = P^2 = (p_1 + p_2)^2 = p_1^2 + p_2^2 + 2p_1 \cdot p_2. \quad (4.23)$$

$$\mathcal{H}_{\text{eff}} = p^0 = \sqrt{\mu^2 + \mathbf{p}^2} = \mu\gamma. \quad (4.25)$$

We can substitute this into the Hamiltonian above, and add back a generic system momentum  $\mathbf{P}$  and obtain

$$\mathcal{H}_{\text{eob}} = \sqrt{M^2(1 + 2\nu(\frac{\mathcal{H}_{\text{eff}}}{\mu} - 1)) + \mathbf{P}^2} \quad (4.26)$$

This Hamiltonian (with  $\mathcal{H}_{\text{eff}}$  given by eq. 4.25) now describes straight line motion in a relativistic framework. The key point is that we can take  $\mathcal{H}_{\text{eff}}$  to be a gravitational test mass Hamiltonian, with some specific deformations and this will match the PN result. Consider a test mass in a Schwarzschild background eq. 2.68, i.e. we consider the case of non-spinning black holes. We can start from the reduced-mass shell expression:

$$g^{\mu\nu}p^\mu p^\nu = -\mu^2 \quad (4.27)$$

where  $g^{\mu\nu}$  is given by eq. 2.68, where we can set  $\theta = \frac{1}{2}\pi$  to fix the orientation<sup>27</sup>. Just as before, the effective Hamiltonian is the 0th component of the momentum, which is implicitly given by eq. 4.27:

$$\mathcal{H}_S = \sqrt{A_S(\mu^2 + A_S p_r^2 + \frac{L^2}{r^2})}, \quad (4.28)$$

where  $A_S = \frac{1}{1-R_S/r}$ ,  $p_r$  and  $L = p_\phi$  are the radial and angular momentum respectively. We then allow deformations from this Hamiltonian defining

$$\mathcal{H}_{\text{eff}}(r, p_r, \phi, L) = \sqrt{A(\mu^2 + A\bar{D}p_r^2 + \frac{L^2}{r^2}) + Q}, \quad (4.29)$$

where  $A, \bar{D}, Q$  are the so-called EOB potentials.  $A, \bar{D}$  correspond to the radial potentials of the effective metric, while  $Q$  contains terms quadratic in the momentum. We can then match these to PN results, by matching observables.

This is more than a neat trick as this Hamiltonian, when well-matched, fits the gold standard NR result up to the merger, well

<sup>27</sup> This is always possible through a change of variables.

out of the PN regime. Crucially, this Hamiltonian framework does not take into account radiation reaction by construction since it is based on conserved energy. These dissipative effects are included after the fact, as friction terms in the EOM. The usual conservative EOM is straightforwardly obtained through Hamilton's equations:

$$\begin{aligned}\dot{r} &= \frac{\partial \mathcal{H}_{\text{eff}}}{\partial p_r}, & \dot{p}_r &= -\frac{\partial \mathcal{H}_{\text{eff}}}{\partial r} \\ \omega = \dot{\phi} &= \frac{\partial \mathcal{H}_{\text{eff}}}{\partial L}, & \dot{L} &= \cancel{-\frac{\partial \mathcal{H}_{\text{eff}}}{\partial \phi}}^0.\end{aligned}\quad (4.30)$$

Note that, as expected, the angular momentum is conserved. We will now spoil this conservation, by adding friction forces, also called radiation reaction forces to the EOMs governing radial and angular momentum dynamics:

$$\begin{aligned}\dot{p}_r &= -\frac{\partial \mathcal{H}_{\text{eff}}}{\partial r} + f_r \\ \dot{L} &= f_\phi.\end{aligned}\quad (4.31)$$

We can now express the energy loss:

$$\dot{E} = \frac{d\mathcal{H}_{\text{eff}}}{dt} = \frac{\partial \mathcal{H}_{\text{eff}}}{\partial r} \dot{r} + \frac{\partial \mathcal{H}_{\text{eff}}}{\partial p_r} \dot{p}_r + \frac{\partial \mathcal{H}_{\text{eff}}}{\partial L} \dot{L} = f_r \dot{r} + f_\phi \omega. \quad (4.32)$$

Such dissipative effects are accounted for by matching to PN results for the angular momentum and energy fluxes, which are given by the so called balance laws:

$$\begin{aligned}\langle \mathcal{F}_L \rangle &= -\langle \dot{L} \rangle = -\langle f_\phi \rangle, \\ \langle \mathcal{F}_E \rangle &= -\langle \dot{E} \rangle = -\langle f_r \dot{r} + f_\phi \omega \rangle.\end{aligned}\quad (4.33)$$

We have denoted averaging with angle brackets. With this we obtain as full of a picture of the inspiral as possible given the inputs, namely PN. The matching can also be done with PM and NR results, further increasing the validity and precision of EOB. Of course, we are still missing the merger and ringdown phases. The merger, or plunge was predicted in Alessandra Buonanno and Damour (2000) to be simply the peak in the waveform. This

waveform is connected smoothly to a ringdown waveform, defined by a sum of quasi-normal modes of the remnant after merger. Thus, one obtains a sort of Frankenstein waveform, pooling resources from PN , PM , NR and BH perturbation theory. Importantly, when well-matched, this waveform is almost as accurate as NR results, but with the added benefit of being able to be computed in a reasonable amount of time. Of course, it is still only as powerful as the inputs to which we match, namely the PN or PM results. We will explore how to compute PM results using amplitudes in the next section.

## 5 Scattering amplitudes and GWs

Whilst we have almost no hope to detect any GW scattering events with current experimental apparatuses, the scattering formalism can still be very useful in the quest for orbital waveforms. Most importantly, while experimental data has not been driving *gravitational* scattering theory, the interplay between *particle* scattering experiment and theory have been an enormously fruitful endeavor for the understanding of fundamental interactions in High Energy Physics (HEP). The back and forth between precise measurements (such as those conducted at the ??) and precise predictions for particle scattering, has pushed the boundaries of the calculations possible. It would be therefore very beneficial if one could apply the large knowledge acquired for small non-gravitationally interacting particles, to large compact orbiting bodies. We will see that it is indeed the case.

Such techniques including EFT, double copy, generalized unitarity, Integration by Parts (IBP) reduction, differential equations methods, developed in the context of perturbative computations of collider observables now find an application to the scattering problem in gravity. A key part of applying tools originally developed for quantum systems to problems in gravity is controlling the classical limit, a historically difficult endeavor. Once scattering data is obtained, one needs to extract the relevant observables. This may be in an effort to compare results from the different methods, in which case gauge and diffeomorphism, and Lorentz invariance are key, since different methods make different gauge and frame choices. It can also be to extract the relevant information for detection and data extraction, in which case the GW waveform is the relevant observable. In the latter case one is mostly interested in the orbital waveform, and thus a map from unbound to bound orbits is necessary.

Multiple formalisms have arisen to map the scattering problem to an orbital one, and the quantum formalism to a classical one. One key way is to map scattering data to a potential (as present in a

Hamiltonian for example). In fact this has been developed as early as the 1970s in (Hiida and Okamura 1972; Iwasaki 1971). Further developments happened in (Neill and Rothstein 2013; Bjerrum-Bohr, Donoghue, and Vanhove 2014; Vaidya 2015; Cachazo and Guevara 2020; Guevara 2019; Damour 2016). Here we will follow the treatment in (Cheung, Rothstein, and Solon Cheung, Rothstein, and Solon 2018), where the conservative part of such scattering amplitudes is matched to an EFT and subsequently mapped to a potential. Recent efforts to add dissipative effects have also been successful (see Kälin, Neef, and Porto 2022). This potential can then be used as input to the EOB formalism. Furthermore, (Kalin and Porto Kalin and Porto 2020a, 2020b), have shown a path forward in directly extending unbound data to bound orbits, obviating the need to use a potential and EOB at all.

Finally, as mentioned in the last chapter, scattering can stand on its own and still be useful without being extended to the orbital case. The PM approximation can be computed from scattering amplitudes in an EFT framework (Kälin and Porto 2020). We can also explore the classical observables possible in scattering scenarios. This has been developed in Kosower, Maybee, and O’Connell (2019) and was extended to local observables<sup>28</sup> such as waveforms in Cristofoli et al. (2022).

In this chapter we will first look at how to use amplitude data to extract orbital and scattering waveforms. We will then look at the KMOC formalism to directly extract observables. Finally, we will look at how to compute amplitudes, and comparing to results such as those in (Kosower, Maybee, and O’Connell 2019; Zvi Bern et al. 2022).

## 5.1 Scattering amplitudes

In QFT, to date the only way to encode scattering data is through a scattering amplitude. Consider a process where one starts out with a set of particles described by a state  $|i\rangle$ , that interact, and yield a final state  $|f\rangle$ . The scattering amplitude is then defined as the probability amplitude for this process to occur. It is simply given by

$$\langle i|S - 1|f\rangle = (2\pi)^4 \delta^{(4)}(p_f - p_i) i \mathcal{A}_{fi}, \quad (5.1)$$

where  $S$  is the S-matrix, encoding asymptotic time evolution,  $S = U_t(-\infty, +\infty)$ . We subtract the identity from it, forming what

<sup>28</sup> local as in not time integrated and thus presenting time dependent dynamics

some call the transfer matrix  $T$ , as we only are interested in processes where something happens (i.e. not everything stays the same). In the above equation we also define the *scattering matrix element*  $\mathcal{A}_{fi}$  (see Srednicki 2007, 76) or *invariant Feynman amplitude* (see Coleman 2018, 214), by factoring out a normalized spacetime delta-function that impose the conservation of momentum on the external particles, and a complex unit. The computation of  $\mathcal{A}$  is made possible by the Lehmann Symanzik and Zimmermann (LSZ) (Lehmann, Symanzik, and Zimmermann 1955; Collins 2019) formalism, which relates  $\mathcal{A}_{fi}$  to time-ordered correlation functions (also known as Green's functions) of the given QFT. Green's functions, can be expanded as an infinite sum of diagrams, weighted by a small coupling. The diagrams encode integrals that are obtained by applying Feynman rules which can be readily extracted from an action describing the theory.

## 5.2 From amplitude to potential

Let us now look at how to map the amplitude to the gravitational potential within the EOB formalism. We will follow the treatment in Cheung, Rothstein, and Solon (2018). We start by defining the effective theory that describes our problem: the orbiting bodies are described by two scalar fields  $f_1$  and  $f_2$  with masses  $m_1$  and  $m_2$  interacting through a long range potential  $V(r)$ . It is the EFT for non-relativistic fields. This EFT is described by the following action

$$S = \int dt \mathcal{L}_{\text{kin}} + \mathcal{L}_{\text{int}}, \quad (5.2)$$

where the kinetic term is given by:

$$\begin{aligned} \mathcal{L}_{\text{kin}} = & \int \tilde{d}^{D-1}\mathbf{k} f_1^*(-\mathbf{k}) \left( i\partial_t - \sqrt{\mathbf{k}^2 + m_1} \right) f_1(\mathbf{k}), \\ & + \int \tilde{d}^{D-1}\mathbf{k} f_2^*(-\mathbf{k}) \left( i\partial_t - \sqrt{\mathbf{k}^2 + m_2} \right) f_2(\mathbf{k}), \end{aligned} \quad (5.3)$$

and the interaction term is given by

$$\mathcal{L}_{\text{int}} = - \int \tilde{d}^{D-1}\mathbf{k} \tilde{d}^{D-1}\mathbf{k}' V(\mathbf{k}, \mathbf{k}') f_1^*(\mathbf{k}') f_2^*(-\mathbf{k}') f_1(\mathbf{k}) f_2(-\mathbf{k}). \quad (5.4)$$

This theory is obtained from the full one by integrating out all the massless force carriers , which are consequently encoded in the potential  $V(\mathbf{k}, \mathbf{k}')$  and taking the non-relativistic limit  $|\mathbf{k}|, |\mathbf{k}'| \ll m_{1,2}$ . In a classical system we assume the particles to be separated by a minimum distance, called impact parameter  $|\mathbf{b}|$ , and consider that their Compton wavelength  $\ell_c \sim \frac{1}{|\mathbf{k}|}, \frac{1}{|\mathbf{k}'|}$ <sup>29</sup> is much smaller than this separation,

$$|\mathbf{b}| \ll \ell_c \simeq \frac{1}{|\mathbf{k}|}, \frac{1}{|\mathbf{k}'|}. \quad (5.5)$$

This ensures that the particles are not interacting quantum mechanically in any significant way. Interestingly this hierarchy of scales can be rewritten as:

$$\mathbf{J} \sim |\mathbf{k} \times \mathbf{b}| \gg 1. \quad (5.6)$$

Thus, any two body classical system has large angular momentum. We can extract the classical part of any quantity by taking the LO contribution in the inverse of the angular momentum, or equivalently:

$$\frac{1}{J} \propto \frac{1}{|b|} \sim |\mathbf{k} - \mathbf{k}'| \propto \frac{1}{\kappa}, \quad (5.7)$$

where  $\kappa$  is the coupling constant, which in the case of gravity is  $\kappa = 4\pi G$ . For the first relation we used that angular momentum is proportional to separation. We then applied the fact that in scattering scenarios the impact parameter is proportional to the inverse of the momentum transfer  $\sim \frac{1}{|\mathbf{q}|}$  where :

$$\mathbf{q} = (0, \mathbf{q}) = (0, \mathbf{k} - \mathbf{k}'). \quad (5.8)$$

The last relation in eq. 5.7 holds due to the virial theorem. Since the potential must encode the coulomb potential  $\kappa/|\mathbf{k} - \mathbf{k}'|^2 \propto J^3$ , it must scale similarly. We thus formulate the following ansatz for the potential:

$$V(\mathbf{k}, \mathbf{k}') = \sum_{n=1}^{\infty} \frac{\kappa^n}{|\mathbf{k} - \mathbf{k}'|^{D-1-n}} c_n \left( \frac{\mathbf{k}^2 + \mathbf{k}'^2}{2} \right) \quad (5.9)$$

<sup>29</sup> Note the use of natural units.

Note that higher order terms in the potential could be formed by any polynomial of momentum invariants  $\mathbf{k}^2$ ,  $\mathbf{k}'^2$ , and  $\mathbf{k} \cdot \mathbf{k}'$ . However, not all combinations of these are independent. The ansatz is chosen such that only the variables  $|\mathbf{k} - \mathbf{k}'|$  and  $\mathbf{k}^2 + \mathbf{k}'^2$  are present, as the others can be reabsorbed by field redefinitions, or vanish on-shell, such as  $\mathbf{k}^2 - \mathbf{k}'^2$ . Note as-well that we work in  $D = 4 - 2\epsilon$  dimensions such that the integrals are dimensionally regulated (see 't Hooft and Veltman 1972). In particular an  $\epsilon$ -power corresponds to a logarithm<sup>30</sup>. Finally, note that in gravity this is precisely a PM expansion!

<sup>30</sup> The third term in the sum is given by  $\kappa^3 \ln(\mathbf{k} - \mathbf{k}')^2 c_3 \left( \frac{1}{2} \mathbf{k}^2 + \mathbf{k}'^2 \right)$ .

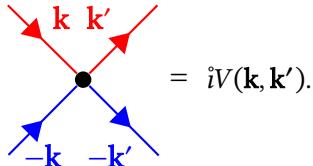
### 5.2.1 EFT amplitude

The first step in establishing determining the potential from a generic scattering amplitude computed in the full theory is to compute the amplitude in the EFT. We first identify the Feynman rules from the action. The kinetic term encodes the propagator:

$$(k_0, \mathbf{k}) = \frac{i}{k_0 - \sqrt{\mathbf{k}^2 + m_{1,2}^2} + i0},$$

Propagator rule

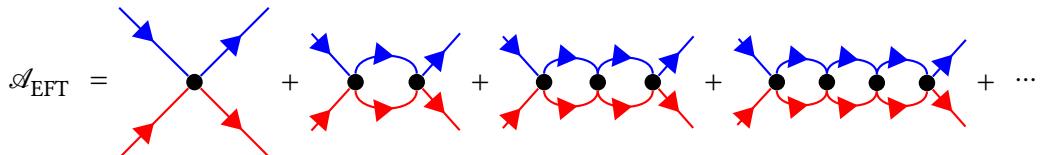
where the  $i0$  is the Feynman prescription for avoiding the poles. The interaction term is encoded in the vertex



Vertex rule

The expression for the EFT amplitude will contain the coefficient functions from eq. 5.9 since they will be present as vertex terms. Particle number must be conserved in the non-relativistic limit<sup>31</sup>, so the amplitude is a sum of bubble diagrams:

<sup>31</sup> As pair production is kinematically forbidden.



## EFT amplitude

We can consequently neatly organize the amplitude into a sum of terms with specific loop counts. We can also equivalently organize it into a sum of terms with specific  $\kappa$  powers. We write,

$$\mathcal{A}_{\text{EFT}} = \sum_{L=0}^{\infty} \mathcal{A}_{\text{EFT}}^{L\text{loop}} = \sum_{i=1}^{\infty} \mathcal{A}_{\text{EFT}}^{(i)}. \quad (5.10)$$

Notably, these partitions do not line up for the EFT since the vertex contains all powers of the coupling. This is in contrast to the full theory where in fact they yield the same partition. We will eventually want to partition over the coupling power to compare the full theory.

Since every cut of the diagrams contributing to  $\mathcal{A}_{\text{EFT}}$  contains two propagators, it is useful to define a 2-body propagator. We can integrate over the energy components of the loop momentum, since the vertex does not have an energy dependence:<sup>32</sup>

$$i\Delta_2(\mathbf{k}) = \int \tilde{d}\omega \frac{i}{\omega - \sqrt{\mathbf{k}^2 + m_1^2}} \frac{i}{E - \omega - \sqrt{\mathbf{k}^2 + m_2^2}} = \frac{i}{E - \sqrt{\mathbf{k}^2 + m_1^2} - \sqrt{\mathbf{k}^2 + m_2^2}} \quad (5.11)$$

where the integral is performed using the residue theorem, closing the contour in either half plane, where in either case a pole is present. We have define  $E = E_1 + E_2$  to be the COM energy of the two initial states:<sup>33</sup>

$$E_i = \sqrt{\mathbf{p}^2 + m_i^2} = \sqrt{\mathbf{p'}^2 + m_i^2}, \quad (5.12)$$

where  $\mathbf{p}$  and  $\mathbf{p'}$  are the initial and final three-momenta of the two states<sup>34</sup> in the COM frame. Finally, we can write the diagram at loop level  $L > 0$  encodes:

$$\mathcal{A}_{\text{EFT}}^{L\text{loop}} = \int \prod_{i=1}^L \tilde{d}^{D-1} \mathbf{k}_i V(\mathbf{p}, \mathbf{k}_1) \Delta_2(\mathbf{k}_1) \cdots \Delta_2(\mathbf{k}_L) V(\mathbf{k}_L, \mathbf{p'}). \quad (5.13)$$

This integral can be performed in the non-relativistic limit, as done by (Bern et al. [Z. Bern et al. 2019](#)) up to 3PM. Up to 2PM, the amplitudes of the EFT is:<sup>35</sup>

<sup>32</sup> The 4-momentum conservation at each vertex means that the energy components of the two propagating momenta must carry along the energy component of the initial state  $E = E_1 + E_2$ . This can be encoded by demanding that  $\sqrt{\mathbf{k}_{\omega_1}^2 + m_1^2} + \sqrt{\mathbf{k}_{\omega_2}^2 + m_2^2} = E$ . Taking the COM frame for the initial momenta means that the 3 momenta of each propagator must cancel i.e.  $\mathbf{k}_1 + \mathbf{k}_2 = 0$

<sup>33</sup> which is equal to the outgoing energy

<sup>34</sup> For example in the initial state one scalar field will have 3-momentum  $\mathbf{p}$ , and the other  $-\mathbf{p}$

<sup>35</sup> Here the re-partitioning mentioned above has been applied and a change in loop momenta:  $\mathbf{k}_i \rightarrow \mathbf{p} + \mathbf{t}_i$ .

$$\begin{aligned}\mathcal{A}_{\text{EFT}}^{(1)} &= -\frac{4\pi G c_1}{\mathbf{q}^2}, \\ \mathcal{A}_{\text{EFT}}^{(2)} &= \pi^2 G^2 \left( -\frac{2c_2}{|\mathbf{q}|} + \frac{1}{E\xi|\mathbf{q}|} \left[ (1-3\xi)c_1^2 + 4\xi^2 E^2 c_1 c'_1 \right] \right) \quad (5.14) \\ &\quad \int \tilde{d}^{D-1} \ell \frac{32E\xi c_1^2}{\ell^2 (\mathbf{q} + \ell)^2 (\ell^2 + 2\mathbf{p}\ell)} \Big),\end{aligned}$$

where  $\xi = \frac{E_1 E_2}{(E_1 + E_2)^2}$  is the reduced energy ratio. The arguments of the coefficient functions  $c_1$  are kept implicit and a prime denotes a derivative with respect to  $\mathbf{p}$ . Additionally, we defined  $\mathbf{q} = \mathbf{p}' - \mathbf{p}$  to be the 3-momentum transfer.

The key point is that the matching procedure requires the EFT amplitude for two-to-two scattering to be equal to the full amplitude at every order in the coupling constant  $\kappa$ , i.e. provided we obtain the amplitude of the full theory, and carefully apply the same limiting procedures, we impose that

$$\mathcal{A}_{\text{EFT}}^{(i)} = \mathcal{A}_{\text{full}}^{(i)} \quad \forall i. \quad (5.15)$$

which fixes the coefficients in the potential ansatz eq. 5.9, and fulfills the promised map from amplitude to potential. Importantly, we must apply the limiting procedure applied above, i.e., the classical and non-relativistic limit, which is summarized by the following hierarchy of scales:

$$m_1, m_2 \ll J|\mathbf{q}| \ll |\mathbf{q}|. \quad (5.16)$$

In the full theory this hierarchy is equivalent to restricting to a specific kinematic regime, following the method of regions. We consider all the possible loop-momentum scalings, first due to the classical limit:

Region	Momentum $\ell = (\omega, \ell)$
hard	$(m, m)$
soft	$( \mathbf{q} ,  \mathbf{q} )$

In this language, the soft region is responsible for the classical limit, since we consider a small  $|\mathbf{q}|$  expansion. We will in fact

explore this expansion in a different light in Section 5.3. Taking the NR limit corresponds to a subregion of the soft region, the so-called potential region:

$$(\omega, \ell) \sim (|\mathbf{u}| |\mathbf{q}|, |\mathbf{q}|), \quad (5.17)$$

where we use that the NR expansion is given by

$$|\mathbf{p}| \ll m \quad (5.18)$$

and thus equivalently <sup>36</sup> by a small relative velocity:

<sup>36</sup> dividing by mass

$$|\mathbf{u}| \ll 1. \quad (5.19)$$

To summarize, in order to obtain the form of the potential, we first compute the amplitude in an EFT, absorbing all force carrying particles into an effective vertex. Making an ansatz for the potential we can compute the amplitude in the non-relativistic limit in terms of unknown coefficients. This same amplitude can be computed in the full theory, which, as a consequence of the limiting procedure, has simplified kinematic dependence. The coefficients of the EFT amplitude are then determined by matching it order by order with the full amplitude.

Unfortunately, the procedure described above, and in fact *any* procedure trying to make contact with EOB by expanding the gravitational potential, intrinsically does not capture the full physics. The crucial point that makes compact binaries interesting is that they are not isolated, energy conserving systems. The energy of the system is lost to the gravitational waves that enable us to detect them. Any conservative dynamics, by construction, cannot capture this loss of energy. In the context of EOB, this radiation reaction was added after the fact, through direct modifications of the EOMs obtained from Hamilton's equations. It is not clear how to incorporate dissipative data from the amplitude in the EOB framework.

One way to not encounter this difficulty is to fully bypass the EOB framework and try to obtain observables directly from the amplitude. This is the subject of the following sections.

## 5.3 KMOC framework

The Kosower Maybee and O’Connell (KMOC) framework (Kosower, Maybee, and O’Connell (2019)) is very general, and is aimed at taking the classical limit of a scattering event in an unspecified theory. We will apply it later on to SQED and gravity.

### 5.3.1 Conventions

First let us define our conventions. Throughout this section we use relativistically natural units, i.e. we do *not* set  $\hbar = 1$ . This will essentially be the small parameter that multiplies the soft momenta that defined the region in Section 5.2. In this section we will further motivate this classical limit in more detail by taking the generally accepted limit that makes quantum physics collapse to classical physics

$$\hbar \rightarrow 0. \quad (5.20)$$

We still retain  $c = 1$ , meaning that, using dimensional analysis we have that  $[L][T]^{-1} = 1$  i.e.  $[L] = [T]$ , length and time are measured in the same units. Correspondingly, energy is measured in units of mass

$$E = mc^2 \implies [E] = [m] = [M], \quad (5.21)$$

and  $\hbar$  has the following units

$$E = \hbar\omega \implies [M] = [\hbar][T]^{-1} \implies [\hbar] = [T][M]. \quad (5.22)$$

Thus, momentum  $p$  has units of  $[p] = [M]$  mass and wavenumber  $[\tilde{p}] = [\frac{p}{\hbar}] = [T]^{-1}$  has units of inverse time. We will denote the Lorentz-invariant phase space measure by  $d\Phi(p)$ <sup>37</sup>:

<sup>37</sup> Repeated integration will be denoted  $d\Phi_n(p_1, \dots, p_n)$

$$\int d\Phi(p) \dots = \int \frac{d^3\mathbf{p}}{2\hbar\omega_{\mathbf{p}}(2\pi)^3} \dots = \int \tilde{d}^4p \tilde{\delta}^{(+)}(p^2 - m^2) \dots, \quad (5.23)$$

where quantities denoted by bars absorb the relevant factors of  $2\pi$ , such that  $\tilde{d}^n p = \frac{d^n p}{(2\pi)^n}$ . Additionally, we have defined  $\hbar\omega_{\mathbf{p}} =$

$\sqrt{\mathbf{p}^2 + m^2}$  to be the on-shell energy and  $\delta^{(+)}$  is the normalized positive energy on-shell delta function:

$$\tilde{\delta}^{(+)}(p^2 - m^2) = (2\pi)\delta(p^2 - m^2)\Theta(p^0) = \tilde{\delta}(p^2 - m^2)\Theta(p^0) \quad (5.24)$$

### 5.3.2 Initial State

With our conventions at hand, let us set the stage for the problem. Imagine we want to scatter two particles, be they massive black holes, or tiny electrons, into each other, with a constant particle number (i.e. a classical two in two out scattering). As mentioned in Section 5.1, in QFT the framework that formalizes scattering of definite particle number states is based on the LSZ reduction formula.

Let us look at it in more detail. The first step is to define the states we want to scatter. Suppose our theory has single particle states  $|p\rangle$ , with mass  $m$ , eigenstates of the momentum:

$$\mathbb{P}^\mu|p\rangle = p^\mu|p\rangle \quad \text{with} \quad p^0 = \hbar\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}. \quad (5.25)$$

These states can be seen as special cases (the plane wave states) of wavepacket states:<sup>38</sup>

$$|f\rangle = \int d\Phi(k) \check{f}(k)|k\rangle, \quad (5.26)$$

where  $\check{f}(k)$  is the momentum space wave-function or, more mathematically, the momentum distribution function. Note that it is almost a Fourier transform, but not quite, as it is performed on the mass shell. If the spacetime ‘wave-function’<sup>39</sup> is given by

$$f(x) = \int d\Phi(k) \check{f}(k)e^{-ik\cdot x}, \quad (5.27)$$

so that  $\check{f}(k)$  is consequently given by

$$\langle k|f\rangle = \check{f}(k) = 2\hbar\omega_{\mathbf{p}} \int d^3\mathbf{x} f(x)e^{ik\cdot x}. \quad (5.28)$$

If we now define the following operators<sup>40</sup> ^[ Note that if we

we work in the Heisenberg picture here

<sup>38</sup> A plane wave state  $|f\rangle = |p\rangle$  would be given by  $\check{f}(k) = 2\hbar\omega_{\mathbf{k}}\tilde{\delta}^{(3)}(\mathbf{k} - \mathbf{p})$

<sup>39</sup> Note that the concept of coordinate space wave-function is ill-defined in interacting theories. However, for such asymptotic states, the wave-function outside the bulk (on the infinite time boundary) is that of a free wave-function.

<sup>40</sup>  $\vec{\partial}_\mu g = f(\partial_\mu g) - (\partial_\mu f)g$

replace  $f(x)$  by a plane wave state  $e^{-ik \cdot x}$  we obtain and define the following:

$$a_{\mathbf{k}}^\dagger(t) = \int d^3\mathbf{x} e^{-ik \cdot x} [\hbar\omega_{\mathbf{k}} f(x) - i\partial_0 f(x)], \quad (5.29)$$

$$a_{\mathbf{k}}(t) = \int d^3\mathbf{x} e^{ik \cdot x} [\hbar\omega_{\mathbf{k}} f(x) + i\partial_0 f(x)], \quad (5.30)$$

]

$$a_f^\dagger(t) = -i \int d^3\mathbf{x} f(x) \overleftrightarrow{\partial}_0 \phi(x), \quad (5.31)$$

$$a_f(t) = i \int d^3\mathbf{x} f^*(x) \overleftrightarrow{\partial}_0 \phi(x), \quad (5.32)$$

where  $\phi(x)$  is the Heisenberg field operator of our theory. While these operators are time dependent, this dependence is not obtained from applying the Heisenberg EOMs<sup>41</sup>, instead it is present through the time dependence of the wave-function  $f(x)$ . It turns out that these operators, suggestively written are the true creation and annihilation operators in the interacting theory.

We can now define a general one particle state  $|f\rangle$  as the result of a creation operator acting on the physical vacuum. Crucially, this only makes sense in the boundary of the bulk, i.e. at asymptotic times  $t \rightarrow \pm\infty$ <sup>42</sup>. Inside the bulk, any thusly created state would in fact not have definite particle number. This definitely makes sense in the scattering problem because if you consider the whole system as a state, it only has definite particle number at asymptotic times, and inside the bulk the scattering happens, and the particle number is not conserved (at least quantum mechanically).

Consequently, we only define asymptotic states, denoting them by an in subscript if they were created at  $t \rightarrow -\infty$  and out if they were created at  $t \rightarrow +\infty$ :

$$|f\rangle_{\text{in}} \stackrel{\text{def}}{=} \lim_{t \rightarrow -\infty} a_f^\dagger(t) |\Omega\rangle = a_{f,\text{in}}^\dagger |\Omega\rangle = \int d\Phi(k) \check{f}(k) \underbrace{a_{\mathbf{k}}^\dagger}_{|k\rangle} |\Omega\rangle \quad (5.33)$$

and the corresponding  $t \rightarrow +\infty$  state where in is replaced with out. It turns out that the extension to multiparticle states is not

<sup>41</sup>  $\frac{dA_H(t)}{dt} = i[H_H, A_H(t)]$

<sup>42</sup> See Coleman (2018) or Collins (2019) for a more in depth discussion of this point

much more complicated. The one thing to demand is that the momentum distributions say  $\check{f}_1, \check{f}_2, \dots$  have no common support, i.e., they are not overlapping. In this case an n-particle in state is defined as:

$$|f_1, \dots, f_n\rangle_{\text{in}} \stackrel{\text{def}}{=} a_{f_1; \text{in}}^\dagger a_{f_2; \text{in}}^\dagger \dots a_{f_n; \text{in}}^\dagger |\Omega\rangle = \int d\Phi_n(k_1, \dots, k_n) \check{f}_1(k_1) \dots \check{f}_n(k_n) |k_1, \dots, k_n\rangle. \quad (5.34)$$

We now have the definitions for asymptotic in and out states for any number of particles. Before defining the setup to scatter two particles, let us look at the meaning of the position space wave-function  $f(x)$  as defined in eq. 5.27. Consider a sharply peaked, compactly supported momentum distribution  $\check{f}(k)$  around a value  $\mathbf{p}_0$  and a characteristic width  $\Delta p$ . One such function could be:

$$f(\mathbf{p}; \mathbf{p}_0, \Delta p) = \begin{cases} N e^{-1/(1-|\mathbf{p}-\mathbf{p}_0|^2/\Delta p^2)} & \text{if } |\mathbf{p}_0 - \mathbf{p}| < \Delta p \\ 0 & \text{if } |\mathbf{p}_0 - \mathbf{p}| \geq \Delta p \end{cases}, \quad (5.35)$$

where  $N$  is a normalization constant. The integrand in eq. 5.27, at large  $t$  or  $\mathbf{x}$ , will be dominated by the stationary phase point  $\mathbf{p}_s$  which is given by:<sup>43</sup>

$$0 = -\frac{\partial}{\partial \mathbf{p}_s} (\hbar \omega_{\mathbf{p}_s} t + \mathbf{p}_s \cdot \mathbf{x}) = -\frac{\mathbf{p}_s t}{\hbar \omega_{\mathbf{p}_s}} + \mathbf{x}. \quad (5.38)$$

In the case of a sharply peaked momentum distribution, the coordinate space wave-function will be largest when the stationary phase point is the same as the peak of the momentum distribution, i.e.,  $\mathbf{p}_s = \mathbf{p}_0$  i.e., substituting  $\mathbf{p}_s$  into eq. 5.38 gives:

$$\mathbf{x} \approx \frac{\mathbf{p}_0 t}{\hbar \omega_{\mathbf{p}_0}} = \mathbf{v}_0 t. \quad (5.39)$$

This is precisely the trajectory of the classical relativistic particle. Now consider the case of two particles with different peaked momentum distributions their trajectories will have different velocities, but the same positions at time 0. Thus, we will shift one of these trajectories by a so called impact parameter  $b^\mu$ <sup>44</sup>, parametrizing the relative separation of the two particles/wave-packets. This can be simply done by multiplying the momentum distribution

<sup>43</sup> This then means that

$$\mathbf{p}_s = \frac{m \mathbf{x} \text{sign}(t)}{\sqrt{t^2 - \mathbf{x}^2}} \quad (5.36)$$

thus

$$\hbar \omega_{\mathbf{p}_s} = \frac{m|t|}{\sqrt{t^2 - \mathbf{x}^2}} \quad (5.37)$$

<sup>44</sup>  $f(x) = \int d\Phi(p) \check{f}(p) e^{-\frac{i}{\hbar} p^\mu x_\mu}$ . Then a shifted, i.e. translated version of  $f(x)$  can be written:

$$\begin{aligned} f(x - x_0) &= \int d\Phi(p) \check{f}(p) e^{-\frac{i}{\hbar} p_\mu (x^\mu - x_0^\mu)} \\ &= \int d\Phi(p) \check{f}(p) e^{\frac{i}{\hbar} p_\mu x_0^\mu} e^{-\frac{i}{\hbar} p_\mu x^\mu} \end{aligned} \quad (5.40)$$

Thus the associated, translated state is:

$$|f\rangle = \int d\Phi(p) \check{f}(p) e^{\frac{i}{\hbar} p_\mu x_0^\mu} |p\rangle \quad (5.41)$$

$\check{f}_1(p)$  by a factor of  $e^{\frac{i}{\hbar}b \cdot p}$ . We take it to be perpendicular to the initial momenta  $p_1, p_2$ . We now can write down the initial state that we are going to study:

$$|in\rangle = \int d\Phi_2(p_1, p_2) \check{f}_1(p_1) \check{f}_2(p_2) e^{\frac{i}{\hbar}b_\mu p_1^\mu} |p_1, p_2\rangle_{in}. \quad (5.42)$$

From now on we will drop the breve and infer from the arguments the type of  $f$  we are dealing with. Observe that by extracting the impact parameter in this way, the wave-functions can be identical in form, and will still be separated as required.

### 5.3.3 Change in observable

The KMOC framework concerns itself with the change of an observable quantity during a scattering event encoded in an operator  $\mathbb{O}$ . For such an observable quantity  $O$ , its change is obtained by evaluating the difference of the expectation value of  $\mathbb{O}$ , between in and out states

$$\Delta O = \langle out | \mathbb{O} | out \rangle - \langle in | \mathbb{O} | in \rangle. \quad (5.43)$$

In quantum mechanics, the out states are related to the in states by the time evolution operator, i.e., the S-matrix,  $|out\rangle = S|in\rangle$ . Thus:

$$\begin{aligned} \Delta O &= \langle in | S^\dagger \mathbb{O} S | in \rangle - \langle in | \mathbb{O} | in \rangle \\ &\stackrel{S^\dagger S = \mathbb{1}}{=} \langle in | S^\dagger [\mathbb{O}, S] | in \rangle \\ &\stackrel{S = \mathbb{1} + iT}{=} \langle in | [\mathbb{O}, \mathbb{1} + iT] | in \rangle - \langle in | iT^\dagger [\mathbb{O}, \mathbb{1} + iT] | in \rangle \quad (5.44) \\ &= \langle in | i[\mathbb{O}, T] | in \rangle + \langle in | T^\dagger [\mathbb{O}, T] | in \rangle \\ &= \Delta O_v + \Delta O_r. \end{aligned}$$

In order, we use the unitarity of the S-matrix, then express the S-matrix as the identity (no actual interaction) and the transfer matrix  $T$ . The commutators are then expanded and the part with the identity vanish (as  $\mathbb{1}$  commutes with everything).

If we put in the definition of our in state (eq. 5.42), we have

$$\Delta O = \int d\Phi_4(p_1, p_2, p'_1, p'_2) f_1(p_1) f_2(p_2) f^*_1(p'_1) f^*_2(p'_2) e^{\frac{ib_\mu p_1^\mu - p'_1^\mu}{\hbar}} [\mathcal{J}_v(O) - \mathcal{J}_r(O)], \quad (5.45)$$

where we defined the real integrand  $\mathcal{J}_r(O)$  and the virtual integrand  $\mathcal{J}_v(O)$  as the following matrix elements<sup>45</sup>

$$\begin{aligned}\mathcal{J}_v(O) &= {}_{\text{in}}\langle p'_1 p'_2 | i[\mathbb{O}, T] | p_1 p_2 \rangle_{\text{in}} \\ \mathcal{J}_r(O) &= {}_{\text{in}}\langle p'_1 p'_2 | T^\dagger [\mathbb{O}, T] | p_1 p_2 \rangle_{\text{in}}.\end{aligned}\quad (5.46)$$

Let us first look at the virtual integrand  $\mathcal{J}_v(O)$ :<sup>46</sup>

$$\begin{aligned}\mathcal{J}_v(O) &= {}_{\text{in}}\langle p'_1 p'_2 | i[\mathbb{O}, T] | p_1 p_2 \rangle_{\text{in}} \\ &= {}_{\text{in}}\langle p'_1 p'_2 | i\mathbb{O}T | p_1 p_2 \rangle_{\text{in}} - {}_{\text{in}}\langle p'_1 p'_2 | iT\mathbb{O} | p_1 p_2 \rangle_{\text{in}} \\ &= iO_{\text{in}'} {}_{\text{in}}\langle p'_1 p'_2 | T | p_1 p_2 \rangle_{\text{in}} - iO_{\text{in}} {}_{\text{in}}\langle p'_1 p'_2 | T | p_1 p_2 \rangle_{\text{in}} \\ &= \underset{p' - p}{i\Delta O} {}_{\text{in}}\langle p'_1 p'_2 | T | p_1 p_2 \rangle_{\text{in}} \\ &= \underset{p' - p}{i\Delta O} \tilde{\delta}^4(p_1 + p_2 - p'_1 - p'_2) \mathcal{A}(p_1, p_2 \rightarrow p'_1, p'_2).\end{aligned}\quad (5.50)$$

Note that the amplitude is from in states to in states. In order to obtain the real integrand  $\mathcal{J}_r(O)$ , we insert a complete set of states:<sup>47</sup>

$$\begin{aligned}\mathcal{J}_r(O) &= {}_{\text{in}}\langle p'_1 p'_2 | T^\dagger [\mathbb{O}, T] | p_1 p_2 \rangle_{\text{in}} \\ &= \sum_X \int d\Phi_{2+|X|}(r_1, r_2, X) {}_{\text{in}}\langle p'_1 p'_2 | T^\dagger | r_1 r_2 X \rangle \langle r_1 r_2 X | [\mathbb{O}, T] | p_1 p_2 \rangle_{\text{in}} \\ &= \sum_X \int d\Phi_{2+|X|}(r_1, r_2, X) \tilde{\delta}^4(p_1 + p_2 - r_1 - r_2 - r_X) \mathcal{A}(p_1, p_2 \rightarrow \underset{r_X - p}{\Delta O} \tilde{\delta}^4(p'_1 + p'_2 - r_1 - r_2 - r_X) \mathcal{A}^*(p'_1, p'_2 \rightarrow r_1, r_2, r_X)),\end{aligned}\quad (5.51)$$

where the  $X$  encodes any number of additional messenger states, and  $r_X$  is the sum of their momenta. For both integrands we can perform some variable changes and eliminate certain Dirac delta functions. We introduce momentum shifts  $q_i = p'_i - p_i$  and then integrate over  $q_2$ , and finally relabel  $q_1 \rightarrow q$ <sup>48</sup>. Thus, we have

$$\begin{aligned}\Delta O_v &= \int d\Phi_2(p_1, p_2) \tilde{\delta}^4 q \tilde{\delta}(2p_1 \cdot q + q^2) \Theta(p_1^0 + q^0) \tilde{\delta}(2p_2 \cdot q - q^2) \Theta(p_2^0 \underset{q}{\Phi(p'_i q^0)} \tilde{\delta}^4 q_i) d\Phi(p_i + q_i) \\ &\quad \times f_1(p_1) f_2(p_2) f_1^*(p_1 + q) f_2^*(p_2 - q) e^{-\frac{i}{\hbar} b_\mu q^\mu} \\ &\quad \times \underset{q}{i\Delta O} \mathcal{A}(p_1, p_2 \rightarrow p_1 + q, p_2 - q),\end{aligned}\quad (5.54)$$

<sup>45</sup> NB: the notation is slightly different in the Zvi Bern et al. (2022) paper

<sup>46</sup> Here we define

$$O_{\text{in}} |p_1 p_2\rangle = \mathbb{O} |p_1 p_2\rangle \quad (5.47)$$

as well as,

$$O'_{\text{in}} \langle p'_1 p'_2 | = \langle p'_1 p'_2 | \mathbb{O} \quad (5.48)$$

and finally

$$\underset{p' - p}{\Delta O} = O'_{\text{in}} - O_{\text{in}} \quad (5.49)$$

<sup>47</sup>  $\sum_X \int d\Phi_{2+|X|}(r_1, r_2, X) |r_1 r_2 X\rangle \langle r_1 r_2 X|$   
Where we could consider the states  $r_1, r_2, X$  to be the final states after the scattering. Note that we always impose having two BHs, at all times, as we have no pair annihilation, which is why the sum always has a two particle dyad. The additional states encoded in  $X$  are all possible additional messenger states (we also have no BH pair production).

<sup>48</sup> Introducing the momentum shifts modifies the measure in the following way:

$$\begin{aligned}&= \tilde{\delta}^4 q_i \tilde{\delta}((p_i + q_i)^2 - m_i^2) \Theta(p_i^0 + q_i^0) \\ &= d\Phi(p_i) \tilde{\delta}((p_i^2 - m_i^2) + 2p_i \cdot q_i + q_i^2)\end{aligned}\quad (5.52)$$

Now since we also have the on-shell enforcing delta function from  $d\Phi(p_i)$  we can rewrite the delta functions:

$$\begin{aligned}&= d\Phi(p_i) \tilde{\delta}((p_i + q_i)^2 - m_i^2) \\ &= d\Phi(p_i) \tilde{\delta}((p_i^2 - m_i^2) + 2p_i \cdot q_i + q_i^2) \\ &= d\Phi(p_i) \tilde{\delta}(2p_i \cdot q_i + q_i^2)\end{aligned}\quad (5.53)$$

Finally we integrate  $q_2$  by solving  $\tilde{\delta}^{(4)}(p_1 + p_2 - p'_1 - p'_2) = \tilde{\delta}^{(4)}(q_1 + q_2)$  and thus we just set  $q_2 = -q_1$

$$\begin{aligned}
\Delta O_r = & \sum_X \int d\Phi_{2+|X|}(r_1, r_2, X) d\Phi_2(p_1, p_2) \tilde{d}^4 q \tilde{\delta}(2p_1 \cdot q + q^2) \Theta(p_1^0 + q^0) \\
& \times \tilde{\delta}(2p_2 \cdot q - q^2) \Theta(p_2^0 - q^0) \\
& \times f_1(p_1) f_2(p_2) f_1^*(p_1 + q) f_2^*(p_2 - q) e^{-\frac{i}{\hbar} b_\mu q^\mu} \\
& \times \frac{\Delta O \tilde{\delta}^{(4)}(p_1 + p_2 - r_1 - r_2 - r_X)}{r_X - p} \\
& \times \mathcal{A}(p_1, p_2 \rightarrow r_1, r_2, r_X) \mathcal{A}^*(p_1 + q, p_2 - q \rightarrow r_1, r_2, r_X).
\end{aligned} \tag{5.55}$$

We have arrived at an integral expression for the change in observable  $\Delta O$ . Luckily for us, we will not need to perform these integrals in the classical limit. We will just have carefully chosen replacement rules for the integrated variables! Let us look at this in more detail now.

### 5.3.4 Classical limit

Since we are concerned with classical observables, we need to explore the classical limit of eq. 5.44, i.e. the limit of  $\hbar \rightarrow 0$ . We first discuss the classical limit of wave-functions. We impose multiple conditions on the wave-functions. The first are those imposed by LSZ reduction. That is,

- Compact support momentum space wave-function
- Peaked around one value of momenta

Furthermore, the classical limit of the wave-functions should make sense, i.e.

1. as  $\hbar \rightarrow 0$  the position and momentum wave-function should approach Dirac delta functions, centered around their classical values.
2. The overlap between the wave-function and its conjugate should be nearly full, since they represent the same particle classically.

Consider for example a non-relativistic wave-function for a particle of mass  $m$ :

$$f(\mathbf{p}) = \exp\left(-\frac{|\mathbf{p}|}{2\hbar m \ell_c / \ell_\omega^2}\right) \stackrel{\hbar=\ell_c m}{=} \exp\left(-\frac{|\mathbf{p}|}{2m^2 \ell_c^2 / \ell_\omega^2}\right), \tag{5.56}$$

where  $\ell_c = \frac{\hbar}{m}$  is the compton wavelength of the particle and  $\ell_\omega$  is a characteristic width. This wave-function, with the proper normalization, grows sharper in the  $\hbar \rightarrow 0$  limit. If we now take the Fourier transform of  $f(\mathbf{p})$  to obtain the position “probability density”, we have:<sup>49</sup>

$$\begin{aligned}\mathcal{F}_{\mathbf{p}}^{-1}[f](\mathbf{x}) &= \int \frac{d\mathbf{p}}{2\pi} \exp\left(-\left(\frac{\mathbf{p}}{A}\right)^2\right) \exp\left(-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}\right) \\ &= \frac{1}{2\pi} \underbrace{\int d\mathbf{p} \exp\left(-\left(\frac{\mathbf{p}}{A} - \frac{i\mathbf{x}A}{2\hbar}\right)^2\right)}_{\sqrt{\pi}A} \exp\left(-\frac{\mathbf{x}^2 A^2}{4\hbar^2}\right) \\ &= \frac{\sqrt{2}A}{2\pi} \exp\left(-\frac{\mathbf{x}^2}{2\ell_\omega^2}\right).\end{aligned}\tag{5.57}$$

<sup>49</sup> *A* absorbs the various constants, with  $A = \sqrt{2m\frac{\ell_c}{\ell_\omega}}$  and  $\mathbf{x}_0$

This elucidates more clearly the meaning of characteristic width, as  $\ell_\omega$  is the standard deviation of the wave-function in position space. Thus, the position-space wave-function grows sharper in the  $\ell_\omega^2 \rightarrow 0$  limit. For both wave-functions to simultaneously grow sharper in the classical limit, we must then have that  $\xi = (\frac{\ell_c}{\ell_\omega})^2 \rightarrow 0$  remembering that the  $\hbar \rightarrow 0$  limit is just given by the  $\ell_c \rightarrow 0$  one. Finally the meaning of classical limit in this context is the  $\xi \rightarrow 0$  limit.

Going back to the general conditions, we want a wave-function  $f_i(p_i)$  such that in the classical limit the momentum reaches its classical value:  $\check{p}_i = m_i \check{u}_i$ , with  $\check{u}_i$  the classical four-velocity of particle  $i$ , normalized to  $\check{u}_i^2 = 1$ . In other words,

$$\langle p_i^\mu \rangle = \int d\Phi(p_i) p_i^\mu |f_i(p_i)|^2 \stackrel{!}{=} m_i \check{u}_i^\mu (1 + \mathcal{O}(\xi^{\beta'})),\tag{5.58}$$

where  $\beta'$  encodes the speed at which the classical value is reached in the  $\xi \rightarrow 0$  limit. The velocity normalization convergence is controlled by  $\beta''$

$$\check{u}_i \cdot u_i = 1 + \mathcal{O}(\xi^{\beta''}),\tag{5.59}$$

Finally wave-function’s spread is controlled by  $\beta$ , and must con-

verge to 0:

$$\begin{aligned}
\sigma^2(p_i)/m_i^2 &= \frac{1}{m_i^2} \langle (p_i - \langle p_i \rangle)^2 \rangle \\
&= \frac{1}{m_i^2} (\langle p_i^2 \rangle - \langle p_i \rangle^2) \\
&= \frac{1}{m_i^2} (m_i^2 - (m_i \check{u}_i (1 + \mathcal{O}(\xi^{\beta'})))^2) \\
&\propto \xi^\beta,
\end{aligned} \tag{5.60}$$

where  $\langle p_i^2 \rangle = m_i^2$  is enforced by the measure  $d\Phi(p)$ .

Additionally, the wave-function should be Lorentz invariant, and naively we would have that  $f(p_i^\mu) = f'(p_i^2)$ . However the integration measure enforces an on-shell condition:  $m_i^2 = p_i^2$ . Thus the wave-function cannot depend on  $p_i^2$ , and we need to introduce at least one additional four vector parameter  $u$ . The simplest dimensionless combination of parameters it then  $\frac{p \cdot u}{m}$ . Of course the wave-function must also depend on  $\xi$  and the simplest form of argument will thus be  $\frac{p \cdot u}{m\xi}$  so that any  $p$  not aligned with  $u$  will be strongly suppressed in the  $\xi \rightarrow 0$  limit.

We now have control over most of the conditions:

- The classical limit is well-defined
- The wave-function spread is controlled
- The arguments of the wave-function are clear

We can write a general wave-function that satisfies the above as:

$$f\left(\frac{p \cdot u}{m} | \check{u}_i; m_i; \beta^{(i)}\right). \tag{5.61}$$

This function can take the form of a Gaussian, or something similar to eq. 5.35. Now there is one final requirement, that concerns the overlap of  $f$  and  $f^*$  must be  $\mathcal{O}(1)$ , equivalently and more precisely:

$$f^*(p+q) \sim f^*(p) \implies f^*(p+q) - f^*(p) \ll 1 \implies q^\mu \frac{\partial}{\partial p^\mu} f^*(p) \ll 1. \tag{5.62}$$

Making explicit the  $\frac{p \cdot u}{m\xi}$  dependence:  $f(p) = \varphi(\frac{p \cdot u}{m\xi})$  for  $\varphi(x)$  a scalar function.

$$\implies \frac{q^\mu u_\mu}{m\xi} \frac{d\varphi^*(x)}{dx} \Big|_{\frac{p \cdot u}{m\xi}} \ll 1. \quad (5.63)$$

Thus we require that for a characteristic value of  $q = q_0$  we have:

$$\frac{q_0 \cdot u}{m\xi} = \bar{q}_0 \cdot u \frac{\ell_\omega^2}{\ell_c} \ll 1 \iff \bar{q}_0 \cdot u \ell_\omega \ll \sqrt{\xi}, \quad (5.64)$$

having we denoted by a bar any quantity that has been rescaled by  $\hbar$ . Thus, a momentum  $p$ , when divided by  $\hbar$  will be written  $\bar{p}$  and called wave-number. We will combine this inequality with ones we obtain from the specific cases of integrations required above.

We now want to examine the classical limit of the integrands of the form eq. 5.54. If we consider just the integration over the initial momenta  $p_i$  and the initial wave-functions with  $\delta(2p_i \cdot q + q^2)$ , the delta function will smear out to a sharply peaked function whose scale is of the same order as the original wave-functions. As  $\xi$  gets smaller, this function will turn back into a Dirac delta function imposed on the  $q$  integration. Let us examine this statement more closely. We are interested in the classical limit of the integrals such as:

$$d(m, \xi, u, q) = \int d\Phi(p) \tilde{\delta}(2p \cdot q + q^2) \Theta(p^0 + q^0) \varphi\left(\frac{p \cdot u}{m\xi}\right) \varphi^*\left(\frac{(p + q) \cdot u}{m\xi}\right). \quad (5.65)$$

This integral must be Lorentz invariant and depends on  $m, \xi, u, q$  thus it must manifestly only depend on the following Lorentz invariants:  $u^2, q^2, u \cdot q, \xi$ . One of these is not actually a variable as we will normalise  $u^2 = 1$ . The rest are not fully dimensionless, but we can render them dimensionless:

$$\begin{aligned} [q^2] &= [\hbar \bar{q}]^2 = [M]^2 \implies [\ell_c \sqrt{-\bar{q}^2}] = \left[\frac{\hbar}{m} \sqrt{-\bar{q}^2}\right] = \frac{[M]}{[M]} = 1, \\ [u \cdot q] &= [M] \implies \left[\frac{u \cdot \bar{q}}{\sqrt{-\bar{q}^2}}\right] = \left[\frac{u \cdot q}{\sqrt{-q^2}}\right] = \frac{[M]}{[M]} = 1, \\ [\xi] &= 1. \end{aligned} \quad (5.66)$$

If we call  $\frac{1}{\sqrt{-\bar{q}^2}} = \ell_s$  a scattering length<sup>50</sup> then our dimensionless ratios become :

$$^{50} [\frac{1}{\sqrt{-\bar{q}^2}}] = [T] = [L] = [\ell_s]$$

$$\frac{\ell_c}{\ell_s} \quad \text{and} \quad \ell_s \bar{q} \cdot u. \quad (5.67)$$

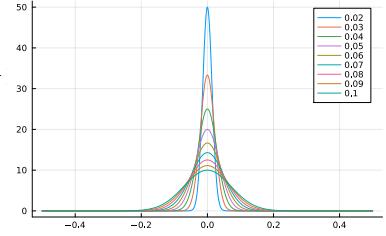
The Dirac delta function can then be rewritten as:

$$\tilde{\delta}(2p \cdot q + q^2) = \tilde{\delta}(2\hbar m u \cdot \bar{q} + \hbar^2 \bar{q}^2) = \frac{1}{\hbar m} \tilde{\delta}(2\bar{q} \cdot u - \frac{\ell_c}{\ell_s^2}) = \frac{\ell_s}{\hbar m} \tilde{\delta}(2\ell_s \bar{q} \cdot u - \frac{\ell_c}{\ell_s}). \quad (5.68)$$

Performing the integration over  $p$  in eq. 5.65 we obtain symbolically:

$$d(m, \xi, u, q) = \text{peaked function imposing that } 2\ell_s \bar{q} \cdot u = \frac{\ell_c}{\ell_s} \text{ with width } \frac{\hbar}{\ell_\omega}. \quad (5.69)$$

Let us discuss the wave-function and the scales from a physical perspective. The characteristic width  $\ell_\omega$  is the particle's position uncertainty,  $\frac{\hbar}{\ell_\omega}$  is the associated momentum uncertainty. In the classical limit the position uncertainty is negligible with respect to minimum distance between the particles  $\ell_s$ :



$$\ell_\omega \ll \ell_s, \quad (5.70)$$

and the momentum uncertainty is negligible with respect to the masses of the particles:

$$\frac{\hbar}{\ell_\omega} \ll m \implies \ell_c \ll \ell_\omega. \quad (5.71)$$

Putting these together we obtain the goldilocks inequality:

$$\ell_c \ll \ell_\omega \ll \ell_s. \quad (5.72)$$

Looking back at the arguments of  $d$  we see that

$$\ell_s \bar{q} \cdot u \gg \ell_\omega \bar{q} \cdot u \sim \sqrt{\xi} = \frac{\ell_c}{\ell_\omega} \gg \frac{\ell_c}{\ell_s}. \quad (5.73)$$

Thus in the classical limit we have that  $d$  collapses to:

$$d(m, \xi, u, q) \propto \tilde{\delta}(\bar{q} \cdot u). \quad (5.74)$$

Thus, the wave-function-weighted on-shell phase-space integration disappears in the classical limit. The resulting condition is that the integration momenta take on their physical values. The sequence that we have gone through should be done for all the integrands of type  $d$ . We will not do this explicitly every time but instead apply the following rules:

1. Just as for the massless transfer momentum any messenger momentum, be it transfer, virtual-loop or real emission, shall become a wave-number:  $k \rightarrow \hbar \bar{k}$ .
2. Replace all couplings with their dimensionless counterparts:  $\kappa \rightarrow \frac{\kappa}{\sqrt{\hbar}}$  (this is only precisely true for SQED and gravity, which are the applications we are interested in).
3. Eliminate all on-shell integrations by approximating  $f(p + \hbar \bar{q})$  by  $f(p)$ .
4. Laurent expand all the integrands in  $\hbar$ .
5. Make the integration momenta take on their physical values:  
 $p_i \rightarrow m_i \check{u}_i$ .

To make this idea explicit we introduce the following notation, meaning that the steps above have been applied

$$\left\langle g(p_1, p_2, \dots) \right\rangle \stackrel{\text{def}}{=} \int d\Phi(p_1) d\Phi(p_2) |f_1(p_1)|^2 |f_2(p_2)|^2 g(p_1, p_2, \dots) \quad (5.75)$$

We can now rewrite eq. 5.44

$$\Delta O = \left\langle \int \overline{d\Psi(q)} \tilde{d}^4 q \tilde{\delta}(2p_1 \cdot q + q^2) \Theta(p_1^0 + q^0) \tilde{\delta}(2p_2 \cdot q - q^2) \Theta(p_2^0 + q^0) e^{-\frac{i}{\hbar} q^\mu b_\mu} (\mathcal{J}'_v(O) + \mathcal{J}'_r(O)) \right\rangle, \quad (5.76)$$

where

$$\begin{aligned}\mathcal{J}'_v(O) &= \frac{i}{q} \Delta O \mathcal{A}(p_1, p_2 \rightarrow p_1 + q, p_2 - q) \\ \mathcal{J}'_r(O) &= \sum_X \int d\Phi_{2+|X|}(r_1, r_2, X) \Delta O \tilde{\delta}^{(4)}(p_1 + p_2 - r_1 - r_2 - r_X) \\ &\quad \times \mathcal{A}(p_1, p_2 \rightarrow r_1, r_2, r_X) \mathcal{A}^*(p_1 + q, p_2 - q \rightarrow r_1, r_2, r_X).\end{aligned}\tag{5.77}$$

Additionally, we make clear the dependence on  $\hbar$  since we want to eventually take the  $\hbar \rightarrow 0$  limit. We apply step 1. above, and change integration variables to  $\bar{q} = \frac{q}{\hbar}$  and absorb the  $\hbar$  dependence into the redefinition of the integrands and the measure:

$$d\Psi(q) = \hbar^2 d\bar{\Psi}(\bar{q}) = \hbar^4 \tilde{d}^4 \bar{q} \frac{1}{\hbar} \tilde{\delta}(2p_1 \cdot \bar{q} + \hbar \bar{q}^2) \Theta(p_1^0 + q^0) \frac{1}{\hbar} \tilde{\delta}(2p_1 \cdot \bar{q} + \hbar \bar{q}^2) \Theta(p_2^0 + q^0),\tag{5.78}$$

$$\begin{aligned}\overline{\mathcal{J}'_v(O)} &= \hbar^2 \mathcal{J}'_v(O), \\ \overline{\mathcal{J}'_r(O)} &= \hbar^2 \mathcal{J}'_r(O).\end{aligned}\tag{5.79}$$

We can finally neatly write:

$$\Delta O = \left\langle \int d\bar{\Psi}(\bar{q}) e^{-i\bar{q}^\mu b_\mu} (\overline{\mathcal{J}'_v(O)} + \overline{\mathcal{J}'_r(O)}) \right\rangle.\tag{5.80}$$

The  $\hbar$  dependence is now entirely contained in the integrands (ignoring the  $\hbar \bar{q}^2$  factors in the delta function). The classical limit of this observable is then simply,

$$\Delta O_{\text{classical}} = \lim_{\hbar \rightarrow 0} \hbar^{\beta_{LO}} \left[ \int d\bar{\Psi}(\bar{q}) (\overline{\mathcal{J}'_v(O)} + \overline{\mathcal{J}'_r(O)}) \right],\tag{5.81}$$

dropping the angle brackets, since these disappear in the classical limit. Here  $\beta_{LO}$  is the LO  $\hbar$ -dependence of the observable. This is so that  $\Delta O_{\text{classical}} \sim \hbar^0$  i.e. classical scaling.

## 5.4 Impulse in KMOC

We can now explore the integrands for a specific observable. Consider the change in momentum, or impulse of particle 1. The KMOC formalisms gives us a way to write this as:

$$\Delta p_1^\mu = \left\langle \int d\bar{\Psi}(\bar{q}) \exp(-i\bar{q}^\mu b_\mu) (\bar{\mathcal{F}}'_v(p_1^\mu) + \bar{\mathcal{F}}'_r(p_1^\mu)) \right\rangle. \quad (5.82)$$

We then have:

$$\begin{aligned} \bar{\mathcal{F}}'_v(p_1^\mu) &= \hbar^2 i q \mathcal{A}(p_1, p_2 \rightarrow p_1 + \hbar \bar{q}, p_2 - \hbar \bar{q}) \\ \bar{\mathcal{F}}'_r(p_1^\mu) &= \hbar^2 d\Phi_{2+|X|}(r_1, r_2, r_X) (r_1^\mu - p_1^\mu) \tilde{\delta}^{(4)}(p_1 + p_2 - r_1 - r_2 - r_X) \\ &\quad \times \mathcal{A}(p_1, p_2 \rightarrow r_1, r_2, r_X) \mathcal{A}^*(p_1 + \hbar \bar{q}, p_2 - \hbar \bar{q} \rightarrow r_1, r_2, r_X). \end{aligned} \quad (5.83)$$

We can extract  $\hbar$  from  $q$  and from the amplitude. In order to determine the  $\hbar$  scaling of  $\mathcal{A}$  we extract each coupling constant  $\kappa$  along with an  $\frac{1}{\sqrt{\hbar}}$ , so that quartic vertices yield a factor of  $\frac{\kappa^2}{\hbar}$  whereas cubic ones yield  $\frac{\kappa}{\sqrt{\hbar}}$ <sup>51</sup>. If we count the number  $V_3$  of all cubic vertices,  $V_4$  the number of quartic vertices, and so on, we have that the number of internal lines is  $I = \frac{1}{2}(\sum_{d=3} dV_d - E)$ . This is because we have  $\sum_{d=3} dV_d$  lines to start with, out of which  $E$  are chosen to be external. The remaining  $(\sum_{d=3} dV_d - E)$  ones are contracted in pairs among themselves to form  $I$  internal lines, yielding the factor of  $\frac{1}{2}$ . In our case we have  $E = 4 + M$  where  $M = |X|$  is the number of messenger particles. Using the argument from loop counting we have that the number of loops of our graph  $L$  is given by:

$$\begin{aligned} L &= I - V + N = \frac{1}{2} \left( \sum_{d=3} d \cdot V_d - 4 - M \right) - \sum_{d=3} V_d + 1 \\ &= \frac{1}{2} \left( \sum_{d=3} (d-2)V_d \right) - 1 - \frac{M}{2}, \end{aligned} \quad (5.84)$$

where  $N$  is the number of connected components ( $= 1$  in our case). Thus, we see that the amount of extracted  $\hbar$ s corresponds directly to the number of loops plus one plus the number of additional messenger particles.<sup>52</sup> We can thus write the amplitude  $\mathcal{A}$  as a

<sup>51</sup> as mentioned in the last section, this is true for gravity and SQED and we will extend this fact to schematically rescale the vertex coupling by  $\hbar^{-\frac{d-2}{2}}$ , for  $d$  the degree of the vertex in question.

<sup>52</sup> The number of extracted couplings being twice that.

sum over reduced  $L$ -loop amplitudes  $\bar{\mathcal{A}}^{(L)}$ :

$$\mathcal{A}(p_1, p_2 \rightarrow r_1, r_2, X) = \sum_{L=0}^{\infty} \left(\frac{\kappa^2}{\hbar}\right)^{(L+1+\frac{|X|}{2})} \bar{\mathcal{A}}^{(L)}(p_1, p_2 \rightarrow r_1, r_2, X). \quad (5.85)$$

Going back to the integrands we have:

$$\begin{aligned} \overline{\mathcal{J}'_V}(p_1^\mu) &= \hbar^3 i \bar{q} \sum_{L=0}^{\infty} \left(\frac{\kappa^2}{\hbar}\right)^{(L+1)} \bar{\mathcal{A}}^{(L)}(p_1, p_2 \rightarrow p_1 + \hbar \bar{q}, p_2 - \hbar \bar{q}) \\ &= i \hbar \bar{q} \sum_{L=0}^{\infty} \kappa^{2(L+1)} \hbar^{(1-L)} \bar{\mathcal{A}}^{(L)}(p_1, p_2 \rightarrow p_1 + \hbar \bar{q}, p_2 - \hbar \bar{q}), \end{aligned} \quad (5.86)$$

as well as the real kernel:<sup>53</sup>

$$\begin{aligned} \overline{\mathcal{J}'_r}(p_1^\mu) &= \hbar^2 d\Phi_{|X|}(r_X) \left[ \prod_{i=1,2} \tilde{d}^4 w_i \tilde{\delta}(2p_i \cdot w_i + w_i^2) \Theta(p_i^0 + w_i^0) \right] \\ &\times w_1^\mu \tilde{\delta}^{(4)}(w_1 + w_2 + r_X) \\ &\times \mathcal{A}(p_1, p_2 \rightarrow p_1 + w_1, p_2 + w_2, r_X) \bar{\mathcal{A}}^*(p_1 + \hbar \bar{q}, p_2 - \hbar \bar{q} \rightarrow p_1 + w_1, p_2 + w_2, r_X) \end{aligned} \quad (5.88)$$

<sup>53</sup> We changed the integration variable from  $r_i$  to  $w_i = r_i - p_i$  thus the measure changes:

$$d\Phi_{2+|X|}(r_1, r_2, X) = d\Phi_{|X|}(r_X) \prod_{i=1,2} \tilde{d}^4 w_i \tilde{\delta}(2p_i \cdot w_i + w_i^2) \Theta(p_i^0 + w_i^0) \quad (5.87)$$

where we used the same reasoning as for the  $p_i$  variable change.

$$\begin{aligned} &= \hbar^2 d\Phi_{|X|}(r_X) \left[ \prod_{i=1,2} \hbar^3 \tilde{d}^4 \bar{w}_i \tilde{\delta}(2p_i \cdot \bar{w}_i + \hbar \bar{w}_i^2) \Theta(p_i^0 + \hbar \bar{w}_i^0) \right] \\ &\times \hbar \bar{w}_1^\mu \hbar^{-4} \tilde{\delta}^{(4)}(\bar{w}_1 + \bar{w}_2 + \bar{r}_X) \\ &\times \sum_{L=0}^{\infty} \left(\frac{\kappa^2}{\hbar}\right)^{(2L+2+|X|)} \bar{\mathcal{A}}^{(L)}(p_1, p_2 \rightarrow p_1 + \hbar \bar{w}_1, p_2 + \hbar \bar{w}_2, r_X) \\ &\times \bar{\mathcal{A}}^{*(L)}(p_1 + \hbar \bar{q}, p_2 - \hbar \bar{q} \rightarrow p_1 + \hbar \bar{w}_1, p_2 + \hbar \bar{w}_2, r_X) \end{aligned} \quad (5.89)$$

$$\begin{aligned} &= d\Phi_{|X|}(r_X) \left[ \prod_{i=1,2} \tilde{d}^4 \bar{w}_i \tilde{\delta}(2p_i \cdot \bar{w}_i + \hbar \bar{w}_i^2) \Theta(p_i^0 + \hbar \bar{w}_i^0) \right] \\ &\times \hbar \bar{w}_1^\mu \tilde{\delta}^{(4)}(\bar{w}_1 + \bar{w}_2 + \bar{r}_X) \\ &\times \sum_{L=0}^{\infty} \kappa^{2(2L+2+|X|)} \hbar^{2-2L-|X|} \bar{\mathcal{A}}^{(L)}(p_1, p_2 \rightarrow p_1 + \hbar \bar{w}_1, p_2 + \hbar \bar{w}_2, r_X) \\ &\times \bar{\mathcal{A}}^{*(L)}(p_1 + \hbar \bar{q}, p_2 - \hbar \bar{q} \rightarrow p_1 + \hbar \bar{w}_1, p_2 + \hbar \bar{w}_2, r_X). \end{aligned} \quad (5.90)$$

Schematically we have

$$\begin{aligned}\overline{\mathcal{J}'_v}(p_1^\mu) &= \sum_{L=0}^{\infty} \mathcal{O}(\kappa^{2(L+1)}), \\ \overline{\mathcal{J}'_r}(p_1^\mu) &= \sum_{L=0}^{\infty} \mathcal{O}(\kappa^{4(L+1)+2|X|})\end{aligned}\quad (5.91)$$

The contributions from the virtual kernel are lower order in the coupling  $\kappa$  for a given loop order. Both kernels contribute together provided that the following equation is verified:

$$L - 1 = 2L' + |X|. \quad (5.92)$$

where  $L$  is the loop count for the virtual kernel and  $L'$ ,  $|X|$  are the real kernel loop count and messenger particle count respectively. Note that for a tree level virtual kernel, the real-kernel match does not exist. The real kernel is only present for  $L > 0$ . When taking the classical limit we will only retain contributions from graphs that cancel the  $\hbar$  divergences in each corresponding kernel. Thus, at the  $L$ -loop level, the amplitude in the virtual kernel must cancel with terms of order

$$\hbar^{1-L+O}, \quad (5.93)$$

where the  $O$  term is the order of  $\hbar$  that is present as a result of the observable. In the case of particle 1's momentum,  $O = 1$ . Similarly, the amplitudes in the real kernel must cancel with:

$$\hbar^{2-2L'-|X|+O}. \quad (5.94)$$

Now we see that the LO contribution<sup>54</sup> to the impulse, which we denote  $\Delta p_1^{\mu,(0)}$  can only be from the virtual kernel at tree level. Thus, we have the following equation,

$$\Delta p_1^{\mu,(0)} = \left\langle \int d\bar{\Psi}(\bar{q}) \exp(-i\bar{q}^\mu b_\mu) \overline{\mathcal{J}'_v}(p_1^\mu)^{(L=0)} \right\rangle. \quad (5.95)$$

And the integrand is given by the tree level 4 point amplitude.

$$\overline{\mathcal{J}'_v}(p_1^\mu)^{L=0} = i\bar{q}^\mu \kappa^2 \hbar^2 \bar{\mathcal{A}}^{(0)}(p_1, p_2 \rightarrow p_1 + \hbar\bar{q}, p_2 - \hbar\bar{q}). \quad (5.96)$$

<sup>54</sup> Here the expansion is in powers of the coupling constant, so even though we want  $\hbar$ s to cancel, the loop order will still affect the order of the contribution through the coupling constant  $\kappa$  and the LO corresponds to  $\kappa$

At Next-to-Leading Order (NLO), i.e.  $\kappa^4$  order, both integrands contribute, as eq. 5.92 can be satisfied for  $L = 1$ ,  $L' = 0$  and  $|X| = 2$ . Thus we have the following equation:

$$\Delta p_1^{\mu,(1)} = \left\langle \int d\bar{\Psi}(\bar{q}) \exp(-i\bar{q}^\mu b_\mu) \left( \overline{\mathcal{F}'_v}(p_1^\mu)^{(L=1)} + \overline{\mathcal{F}'_r}(p_1^\mu)^{(L'=0)} \right) \right\rangle. \quad (5.97)$$

The virtual integrand is now given by the 1-loop level amplitude, and the real integrand is given by the square of a the tree level amplitude. This process can go on indefinitely, and is independent of the type of observable and the theory. Here we considered the change of momentum a particle, which for a black hole very far away would be very difficult to measure. However, we can also consider an observable such as the four-momentum of the radiated particles, or more precisely its expectation value. Of course, the operator corresponding to this observable gives zero when acting on the initial momentum states, and only gives a non-zero result when acting on the messenger states. Thus, for this observable only the real integrand, starting with  $|X| = 1$  will contribute, the LO contribution being given by what is essentially the unitarity cut of a two loop amplitude. We see that regardless of observable, the objects that are needed are the amplitudes.

For each loop level, many diagrams can contribute, but the classical limit enforces that they must cancel the  $\hbar$  orders given by eq. 5.94 and eq. 5.93. The cancellation order is dependent on the considered observable and filters the contributing diagrams. It can also be reformulated in the language of the method of regions.

To see the whole machinery in action, let us take SQED as an example theory, that shows the relevant subtleties of the formalism.

## 5.5 SQED amplitudes

We want to couple a set of massive scalar fields to electromagnetism. We will use the minimal coupling prescription to ensure that the resulting Lagrangian exhibits the required gauge symmetry. The minimal coupling prescription works provided that our free mass Lagrangian admits a conserved current. Our free mass

Lagrangian for two complex scalar fields with different masses is:

$$\mathcal{L}_m = \sum_{i=1}^2 \partial_\mu \phi_i^\dagger \partial^\mu \phi_i - m_i^2 \phi_i^\dagger \phi_i. \quad (5.98)$$

Notice that under the following transformation:

$$\begin{aligned} \phi_i(x) &\rightarrow e^{-iQ_i\lambda} \phi_i(x), \\ \phi_i^\dagger(x) &\rightarrow e^{iQ_i\lambda} \phi_i^\dagger(x), \end{aligned} \quad (5.99)$$

or infinitesimally:

$$\begin{aligned} \delta\phi_i &= -i\phi_i Q_i \delta\lambda, \\ \delta\phi_i^\dagger &= i\phi_i^\dagger Q_i \delta\lambda, \end{aligned} \quad (5.100)$$

the above Lagrangian is unchanged. We will identify  $Q_i$  with the charge, in units of  $e$ , of each particle. If we upgrade the parameter  $\lambda$  to a spacetime function  $\lambda(x)$ , the Lagrangian is not invariant anymore and invariance is restored if we replace all the derivatives  $\partial_\mu$  with the gauge-covariant derivatives  $D_\mu$ :

$$D_\mu = \partial_\mu + ieQ_i A_\mu, \quad (5.101)$$

i.e., if along with the transformation above, we perform a gauge transformation of the photon field:

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu \lambda \quad (5.102)$$

or infinitesimally

$$\delta A_\mu = \frac{1}{e}\partial_\mu \delta\lambda, \quad (5.103)$$

then the whole Lagrangian:

$$\mathcal{L}_{ED} + \mathcal{L}_m(\phi_i, \partial_\mu \phi_i) \rightarrow \mathcal{L}_{ED} + \mathcal{L}_m(\phi_i, D_\mu \phi_i), \quad (5.104)$$

is invariant under the above defined gauge transformation. The final Lagrangian is:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \sum_{i=1}^2 \left[ (D_\mu\phi_i)^\dagger (D^\mu\phi_i) - m_i^2\phi_i^\dagger\phi_i \right], \quad (5.105)$$

with  $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ . Expanding and then integrating by parts:

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) + \sum_{i=1}^2 \left[ (\partial_\mu - ieQ_i A_\mu)\phi_i^\dagger (\partial_\mu + ieQ_i A_\mu)\phi_i - m_i^2\phi_i^\dagger\phi_i \right] \\ &= \underbrace{\frac{1}{2}A_\mu [\eta^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu] A_\nu}_{\mathcal{L}_0} + \underbrace{\sum_{i=1}^2 -\phi_i^\dagger (\partial^2 + m_i^2)\phi_i}_{\mathcal{L}'} + \underbrace{\sum_{i=1}^2 q_i^2 e^2 A_\mu A^\mu \phi_i^\dagger \phi_i - ieQ_i A_\mu (\phi_i^\dagger \partial^\mu \phi_i - (\partial^\mu \phi_i^\dagger) \phi_i)}_{\mathcal{L}'} \end{aligned} \quad (5.106)$$

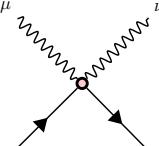
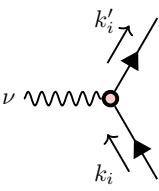
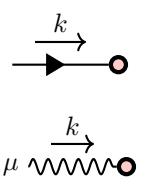
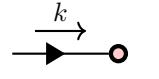
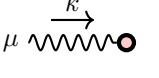
Since we have a massless photon, and still have gauge freedom we have to implement gauge fixing in the usual way:

$$\begin{aligned} \mathcal{L}_{\text{eff}} &= \mathcal{L} + \overbrace{\frac{-1}{2\xi}(\partial_\mu A^\mu)^2}^{\mathcal{L}_{\text{GF}}} \\ &= \frac{1}{2}A_\mu \left[ g^{\mu\nu}\partial^2 - \left(1 - \frac{1}{\xi}\right)\partial^\mu\partial^\nu \right] A_\nu + \sum_{i=1}^2 -\phi_i^\dagger \underbrace{(\partial^2 + m_i^2)}_{\tilde{\Delta}_F(q^2) = \frac{i}{p^2 - m^2 + i\varepsilon}} \phi_i \\ &\quad \rightarrow \tilde{D}_\xi^{\mu\nu}(k) = \frac{-i}{k^2 + i\varepsilon} \left[ g_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^2} \right] \\ &+ \sum_{i=1}^2 \underbrace{Q_i^2 e^2 \eta^{\mu\nu}}_{\times 2i \rightarrow 4-\text{vertex}} A_\mu A_\nu \phi_i^\dagger \phi_i - \underbrace{ieQ_i A_\mu (\phi_i^\dagger \partial^\mu \phi_i - (\partial^\mu \phi_i^\dagger) \phi_i)}_{\times i \rightarrow 3-\text{vertex}}. \end{aligned} \quad (5.107)$$

In the last lines we identify the terms contributing to the Feynman rules. They are given in Table 5.2

Tbl 5.2: Feynman rules for SQED

For every	Write
$\mu \xrightarrow{\ell} \nu$ internal photon line	$\tilde{D}_\xi^{\mu\nu}(\ell) = \frac{-i}{\ell^2 + i\varepsilon} \left[ \eta_{\mu\nu} - (1 - \xi) \frac{\ell_\mu \ell_\nu}{\ell^2} \right]$
$\xrightarrow{k}$ internal scalar	$\tilde{\Delta}_F(k^2) = \frac{i}{k^2 - m^2 + i\varepsilon}$

For every	Write
	$\phi_i \phi_i^\dagger A_\mu A_\nu$ vertex
	$2iQ_i^2 e^2 \eta^{\mu\nu}$
	$\phi_i(k_i) \phi_i^\dagger(k'_i) A_\mu$ vertex
	External scalar $\times 1$
	$\left\{ \begin{array}{l} \text{incoming} \\ \text{outgoing} \end{array} \right\}_{\text{photon}} \times \left\{ \begin{array}{l} \varepsilon_\mu \\ \varepsilon_\mu^* \end{array} \right\}, \text{with}$ $\varepsilon \cdot k = 0, \varepsilon' \cdot k' = 0$

where we consider only incoming momenta and the arrows denote incoming or outgoing particles. For the photon, we will take the Feynmann gauge, setting  $\xi = 1$  and thus  $\tilde{D}_1^{\mu\nu}(\ell) = \tilde{D}^{\mu\nu}(\ell) = \frac{-i\eta^{\mu\nu}}{\ell^2 + i\varepsilon}$ .

### 5.5.1 Expansions and simplifications

Notice that if we rescale all photon momenta (taking them to be the loop momenta) by  $\hbar$  or equivalently take them to be in the soft region:  $|k| \sim |q| \ll |p|$ , where  $p$  is the external momenta, then the photon propagator scales homogeneously in  $\hbar$ , typically:

$$\frac{1}{(\ell - q)^2} = \frac{1}{\hbar^2} \frac{1}{(\bar{\ell} - \bar{q})^2}, \quad (5.108)$$

thus contributes  $\mathcal{O}(\hbar^{-2})$  to the overall diagram. The matter propagator on the other hand, has inhomogeneous  $\hbar$  scaling, as we do not rescale the external momenta, (or equivalently the external momenta are not restricted by the regions). Note that we will not consider internal mass loops as the soft limit means that massive

pair production is forbidden, thus massive propagators necessarily route an external momentum. We can nonetheless expand a generic massive propagator in the soft limit:

$$\begin{aligned} \frac{1}{(\ell - p_i)^2 - m_i^2} &= \frac{1}{\ell^2 - 2\ell \cdot p_i + \cancel{p}_i^2 - \cancel{m}_i^2} = \frac{1}{\hbar^2 \bar{\ell}^2 - 2\hbar \bar{\ell} \cdot p_i} \\ &= -\frac{1}{\hbar} \frac{1}{2\bar{\ell} \cdot p_i} (1 + \hbar \frac{\bar{\ell}^2}{2\bar{\ell} \cdot p_i} + \hbar^2 \frac{\bar{\ell}^4}{(2\bar{\ell} \cdot p_i)^2} + \dots). \end{aligned} \quad (5.109)$$

In other contexts one may say that the matter propagator has eikonalized (it is now linear in  $\ell$ ). Before we compute the amplitudes let us set up some useful kinematic identities and variables. If we consider the two-to-two particle scattering, taking an all outgoing momentum convention we have the following masses:

$$m_1^2 = p_4^2 = p_1^2, \quad m_2^2 = p_2^2 = p_3^2, \quad (5.110)$$

and Mandelstam variables:

$$s = (p_1 + p_2)^2, \quad t = q^2 = (p_1 + p_4)^2, \quad u = (p_1 + p_3)^2, \quad (5.111)$$

subject to the usual equation:

$$s + t + u = 2(m_1^2 + m_2^2). \quad (5.112)$$

We can also change the external momentum variables to ones more amenable to the soft limit, namely:

$$p_1 = -\left(\tilde{p}_1 - \frac{q}{2}\right), p_2 = -\left(\tilde{p}_2 + \frac{q}{2}\right), p_3 = \left(\tilde{p}_2 - \frac{q}{2}\right), p_4 = \left(\tilde{p}_1 + \frac{q}{2}\right). \quad (5.113)$$

The new momentum variables  $\tilde{p}_i$  are crucially orthogonal to momentum transfer  $q$ :<sup>55</sup>

<sup>55</sup> we use eq. 5.110

$$\tilde{p}_i \cdot q = 0, \quad (5.114)$$

and the physical scattering region, given by  $s > (m_1 + m_2)^2$  and  $q^2 < 0$  is given by the same formulas:

$$s = \left( -\left(\tilde{p}_1 - \frac{q}{2}\right) - \left(\tilde{p}_2 + \frac{q}{2}\right) \right)^2 = (\tilde{p}_1 + \tilde{p}_2)^2 \quad (5.115)$$

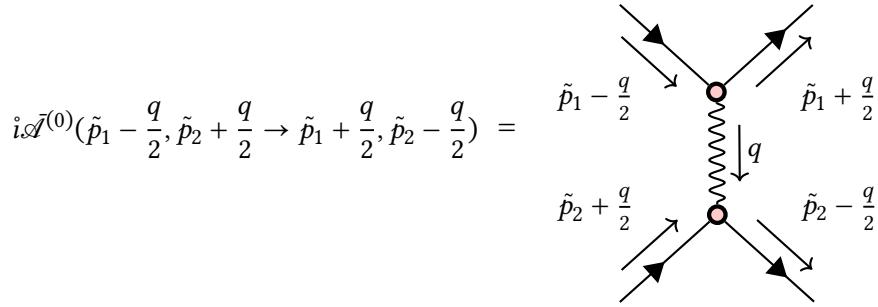
and

$$t = \left( -\left(\tilde{p}_1 - \frac{q}{2}\right) + \tilde{p}_1 + \frac{q}{2} \right)^2 = q^2. \quad (5.116)$$

With all the ingredients in place, we can now go on to compute the amplitudes.

### 5.5.2 Tree level

We start with the tree level amplitude, the only LO contribution in for example eq. 5.95. The only possible diagram we can build with four external scalar legs, and vertices as defined in Table 5.2, is the following tree:



Tree

The amplitude is read off diagram and using Feynman rules for SQED we have:

$$\mathcal{A}^{(0)} = i\tilde{D}^{\mu\nu}(q) \cdot e^2 Q_1 Q_2 2\tilde{p}_1^\mu 2\tilde{p}_2^\nu = \frac{4e^2 Q_1 Q_2 \tilde{p}_1 \cdot \tilde{p}_2}{q^2}, \quad (5.117)$$

using the Mandelstahm invariants described above we can write this as:

$$\mathcal{A}^{(0)} = e^2 Q_1 Q_2 \frac{4m_1 m_2 \sigma}{\hbar^2 \tilde{q}^2}, \quad (5.118)$$

where  $\sigma$  is the relativistic factor of particle 1 in the rest frame of particle 2:

$$\sigma = \frac{s - m_1^2 - m_2^2}{2m_1 m_2} = \frac{p_1 \cdot p_2}{m_1 m_2} = \frac{\tilde{p}_1 \cdot \tilde{p}_2}{m_1 m_2} + \mathcal{O}(\hbar). \quad (5.119)$$

We now input the reduced version  $\tilde{\mathcal{A}}^{(0)} e^2 = \mathcal{A}^{(0)}$  of the amplitude, and take the  $\hbar \rightarrow 0$  limit of eq. 5.95. We can safely take the  $\hbar \rightarrow 0$  limit as the integrand contains no terms singular in  $\hbar$  (the  $\frac{1}{\hbar^2}$  is cancelled by the  $\hbar^2$  pre-factor). Notice that the integration measure eq. 5.76 simplifies in the classical limit:<sup>56</sup>

$$\lim_{\hbar \rightarrow 0} d\bar{\Psi}(\bar{q}) = \tilde{d}^4 \bar{q} \tilde{\delta}(2m_1 \check{u}_1 \cdot \bar{q}) \tilde{\delta}(2m_2 \check{u}_2 \cdot \bar{q}). \quad (5.121)$$

The integrand corresponding to the LO contribution to the impulse is then:

$$\Delta p_1^{\mu, (0)} = 4e^2 Q_1 Q_2 m_1 m_2 \sigma \int d\tilde{q} e^{-i\tilde{q} \cdot b} \tilde{q}^\mu \frac{\hbar^2}{\hbar^2 \tilde{q}^2} \tilde{\delta}(2m_1 \check{u}_1 \cdot \tilde{q}) \tilde{\delta}(2m_2 \check{u}_2 \cdot \tilde{q}). \quad (5.122)$$

This can be analytically computed to find a closed form for the LO impulse.

$$\Delta p_1^{\mu, (0)} = -\frac{e^2 Q_1 Q_2}{2\pi} \frac{\gamma}{\sqrt{\gamma^2 - 1}} \frac{b^\mu}{b^2} \quad (5.123)$$

### 5.5.3 One loop

Already at one loop the number of possible diagrams increases dramatically from 1 to 13. One can nonetheless enumerate them. As we increase loop count, the number of diagrams increases exponentially. It is therefore useful to have a systematic, programmatic way to handle these diagrams, and while we are at it, also to compute the amplitudes. Note that all the code needed to do these computations can be found on the repository for this book: <https://github.com/lcnhb/GWAmplitudes>.

The first tool we will use is QGRAF Nogueira (1993), a Fortran based program that can generate Feynman diagrams for a given set of

<sup>56</sup> Compare to

$$\begin{aligned} d\bar{\Psi}(\bar{q}) &= \tilde{d}^4 \bar{q} \tilde{\delta}(2p_1 \cdot \bar{q} + \hbar \bar{q}^2) \Theta(p_1^0 + q^0) \\ &\quad \tilde{\delta}(2p_1 \cdot \bar{q} + \hbar \bar{q}^2) \Theta(p_2^0 + q^0) \end{aligned} \quad (5.120)$$

The theta functions cancel as  $q^0 \rightarrow 0$  and  $p_i$  becomes classical.

vertices and external legs. The first step is to define the theory we want to derive the diagrams from. This is essentially writing down the Feynman rules, without actually specifying what the graphical objects are. For SQED the model file is short:

### SQEDmodel

```
* propagators
[Phi1,Phi1c,+]
[Phi2,Phi2c,+]
[Photon,Photon,+]

* vertices
[Phi1c,Photon,Phi1]
[Phi2c,Photon,Phi2]
[Phi1c,Photon,Photon,Phi1]
[Phi2c,Photon,Photon,Phi2]
```

This file also defines the names of the fields in question. We can then give QGRAF additional options, such as the number of external legs, and the number of loops, written in a `qgraf.dat` file that can look like this:

### qgraf.dat

```
output= '{{output filename}}' ;
style= './styles/julia.sty' ;

model= './qgraf/SQEDmodel';

in= Phi1, Phi2;
out= Phi1,Phi2;
loops= 1;
loop_momentum= k;

options= notadpole,nosnail,onshell ;
```

where we specify what process we are interested in (two-to-two scattering). We also specify that we want to compute the one loop amplitude, and that we want to use the `julia.sty` style file. The `julia.sty` file is a style file that is used to generate the output in a format that can be read by a Julia ([Bezanson et al. 2015](#)) package built to process and visualise the outputted graphs. This package is called (maybe too broadly) `QFT.jl`, and is hosted on [github](#). Let

us see it in action. We can run QGRAF from the command line, and it will generate a file called `{output filename}` that contains the output of the program. We can then use this in a Julia file to visualize the diagrams.

```

using Images
using QFT
using QFT.Diagrams
using QFT.FieldGraphs
using QFT.Fields
import Catlab.Graphics.Graphviz: pprint
include("SQED.jl")
diags=include("QGRAFout/julia/1lSQED.jl")

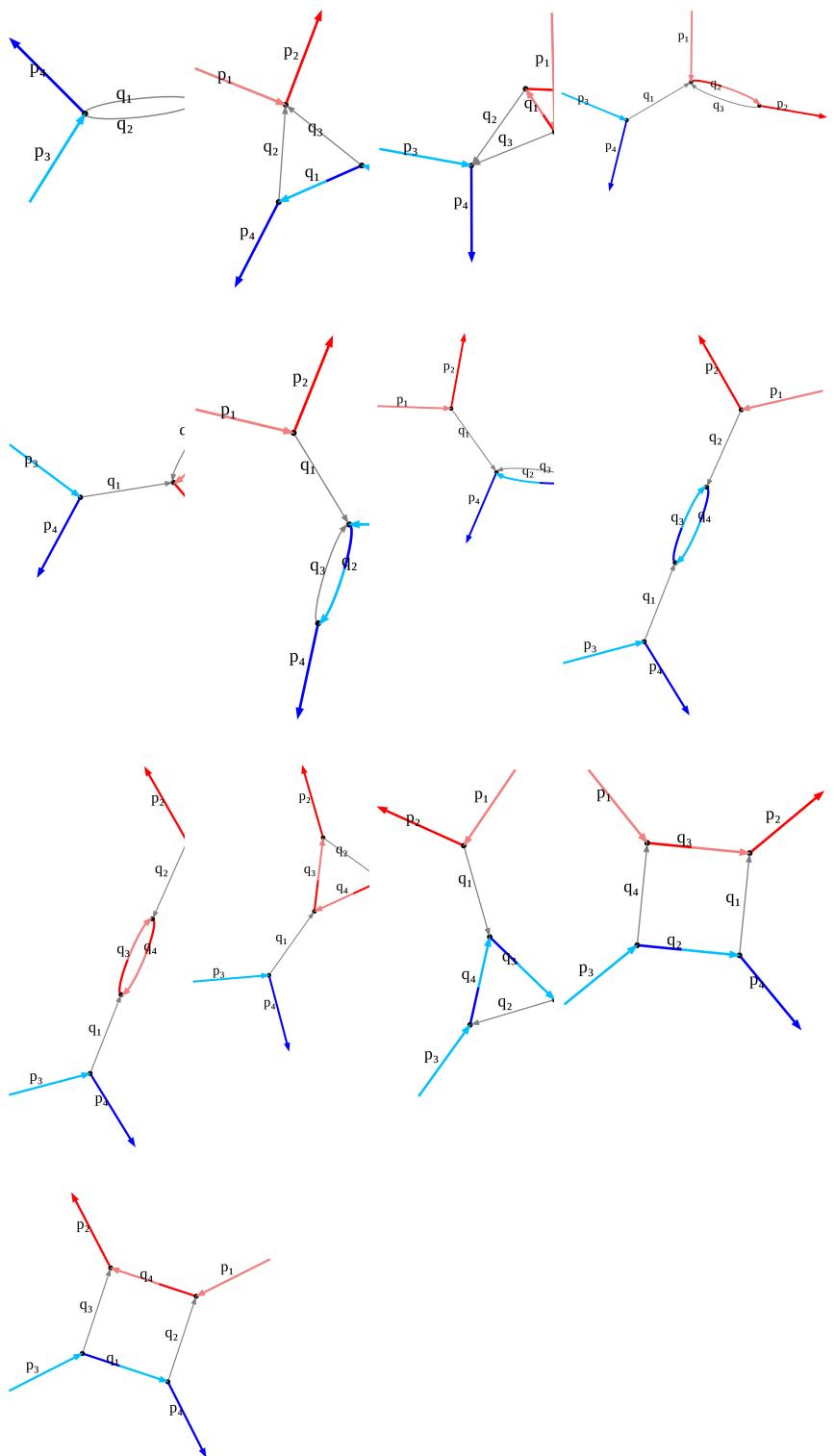
qDiags=[qDiagram(;diag...) for diag in diags]
grafs=(x -> x.g).(qDiags)
println("::: {#fig-oneoopsqed layout=ncol=4 .column-body-right }")
for (i,graf) in enumerate(grafs)
    println("``{dot}")
    println("//| fig-width: 1.5")
    pprint(stdout,to_graphviz(graf,graph_attrs=Dict("layout"=>"neato")))
    println("``")
end
println("all one loop graphs in (?:sqed)")
println(":::")
nothing

```

We notice a few things about the diagrams in Figure ?? above. Not all the diagrams contribute to the classical limit. Clearly the diagrams with internal matter loops cannot be allowed when the momenta incoming to the loop is soft<sup>57</sup>. In the classical limit this momentum will go to 0, and since the scalars are massive, they need momentum greater than their mass to be pair-produced, thus these diagrams do not contribute. Another class of diagrams that does not contribute to the classical limit are the ones with photon loops that start and end on the same matter line. These do not contribute because they are scaleless. An integral is scaleless when its integrand  $\mathcal{I}_{\text{scaleless}}(\{\ell_i\})$  scales homogeneously under any rescaling of the loop momentum  $\ell_i$ :

$$\mathcal{I}_{\text{scaleless}}(\{\ell_i\}) \xrightarrow{\ell_i \rightarrow \lambda \ell_i} \lambda^\eta \mathcal{I}_{\text{scaleless}}(\{\ell_i\}) \quad (5.124)$$

<sup>57</sup> scales like  $\hbar$



Thus, any diagram that has a basis of loops that does not contain both scalar fields (i.e. does not cross over) cannot have its scale set by the momentum transfer, and is thus scaleless. Dimensionally regularized scaleless integrals vanish, thus we can also discard these diagrams. We can easily implement such a filter programmatically, by checking that if cycles are present they touch either both types of scalars, or neither. To compute the cycle basis we use an algorithm derived from Paton (1969), implemented in the [DirectedHalfEdgeGraphs.jl](#) package.

```

using Logging
Logging.disable_logging(Logging.Info)
function scaleless(g::AbstractFieldGraph)
    cycles =cycle_basis(g)
    isscaleless=false
    for c in cycles
        hs= collect(Iterators.flatten(half_edges.(Ref(g),c)))
        fields=unique!(typeof.(field.(Ref(g),hs)))
        if Bool(Phi1 in fields) ∨ Bool(Phi2 in fields)
            isscaleless=true
        end
    end
    return isscaleless
end

noscalelessdiags=qDiags[.!scaleless.(grafs)]
classicaldiags=noscalelessdiags[2:end]
for (i,d) in enumerate(classicaldiags)
    d.ID=i
end

println("::: {#fig-onefilterloopsqed layout-ncol=5 .column-body-right }")
for (i,graf) in enumerate(grafs[.!scaleless.(grafs)])
    println("``{dot}")
    println("//| fig-width: 1")
    pprint(stdout,to_graphviz(graf,graph_attr=Dict("layout"=>"neato")))
    println("``")
end
println("one loop graphs that contribute to the classical limit in (?:sqed)")
println(":::")
nothing

```

There is one final simplification we can do. The first diagram in

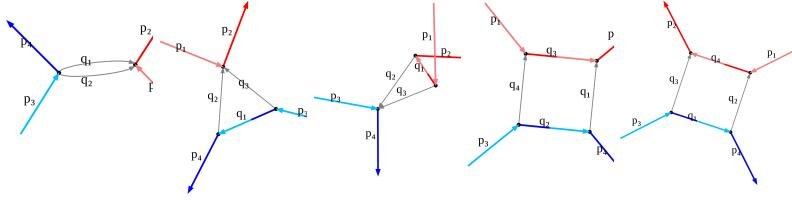


Figure ?? has two photon propagators, and two quartic vertices, which give it the name: “double-seagull”. The photon propagators have homogeneous  $\hbar^{-2}$  scaling (see eq. 5.108), the vertices only bring constants, and the loop integration has  $\hbar^4$  scaling. Collecting all scalings this means that the reduced amplitude has  $\hbar^0$  scaling. However, in the virtual integrand the reduced amplitude has to cancel with the transfer momentum’s  $\hbar$  scale (see eq. 5.90). Thus, the double seagull diagram vanishes in the classical limit.

We are left with only 4 diagrams, the two triangle diagrams, the box, and cross box, which we can now compute. At this point, we can do this manually, but again, putting in place a programmatic framework means that subsequent loops can be treated systematically<sup>58</sup>. We will also extract one of these integrands by hand, for completeness and comparison.

The first step is to define the Feynman rules in a computer readable format. We will use `Julia` to apply these rules at every vertex and every edge, but for the actual computation we will use `FORM` (Vermaseren 2000; Kuipers et al. 2013), which can automatically contract indices and performs expression manipulations in an optimized way. The Feynman rules implementation in the package developed for this thesis make use of `Julias` defining paradigm of multiple dispatch. This allows us to define the rules for each vertex and each edge in a very compact way. We can define the rules for SQED as follows:

```
function feynmanRule(mime::MIME"text/FORM",p,f1::Photon,f2::Photon)
    momen = repr(mime,p.symbol)
    idx1 = repr(mime,index(f1))
    idx2 = repr(mime,index(f2))
    "_i* prop(-1,0,$momen)*d_($idx1,$idx2)"
end

@symmetric function feynmanRule(mime::MIME"text/FORM",p,[f1::Phi1,f2::Phi1c])
```

<sup>58</sup> at least at the un-integrated level

```

momen = repr(mime,p.symbol)
"_i* prop(-1,1,$momen)"
end

@symmetric function feynmanRule(mime::MIME"text/FORM",p,[f1::Phi2,f2::Phi2c])
momen = repr(mime,p.symbol)

"_i* prop(-1,2,$momen)"
end

@symmetric function feynmanRule(mime::MIME"text/FORM",[(pφ1,a)::Tuple{Any,ScalarField{S}},
(pφ2,b)::Tuple{Any,AdjointField{ScalarField{S}}},(pγ,γ)::Tuple{Any,Photon}]) where {S}
p1 = repr(mime,pφ1(index(γ)))
p2 = repr(mime,pφ2(index(γ)))
q = repr(mime,charge(a))
"_i*e*q*($p2-$p1)"
end

@symmetric function feynmanRule(mime::MIME"text/FORM",[(pγ1,γ1)::Tuple{Any,Photon},
(pγ2,γ2)::Tuple{Any,Photon},(pφ2,b)::Tuple{Any,AdjointField{ScalarField{S}}},(pφ1,a)::Tuple{Any,ScalarField{S}}]) where {S}
mu = repr(mime,index(γ1))
nu = repr(mime,index(γ2))
q = repr(mime,charge(a))
"2*i*q^2*e^2*d_($mu,$nu)"
end

```

where we define `prop(n,i,p)` as representing a generic Feynman propagator, which we will denote:

$$D_-(n, i, p) = \text{prop}(n, i, p) = (p^2 - m_i^2 + i\epsilon)^n \quad \text{thus} \quad \text{prop}(-1, i, p) = \frac{1}{p^2 - m_i^2 + i\epsilon}. \quad (5.125)$$

We can now apply these Feynman rules to the graphs we have left, obtaining FORM readable integrands. For example, the box diagram gives:

```

```{julia}
toform(stdout, classicaldiags[3])
```

```

```

*--#[d3l1:

L [d3l1|o.3.4|i.3.4|i|o|] = 1*i_* prop(-1,0,q1)*d_(nu1,nu2)*i_* prop(-
1,1,q2)*i_* prop(-1,2,q3)*i_* prop(-1,0,q4)*d_(nu7,nu8)*i_*e*qch1*(q2(nu1)-
p4(nu1))*i_*e*qch2*(q3(nu2)-p2(nu2))*i_*e*qch1*(p3(nu7)-
q2(nu7))*i_*e*qch2*(p1(nu8)-q3(nu8));
#procedure momentumRouting

Id q2 = p3 + -q4;
Id p1 = p2 + p4 + -p3;
Id q1 = p3 + -p4 + -q4;
Id q3 = p2 + p4 + q4 + -p3;

Id q4= l1;
#endprocedure
*--#]d3l1:

```

Notice that we also define the momentum routing by imposing the replacement rules, one per vertex. This can be done in many ways but to ensure proper scaling we set the loop momenta to be homogeneous on the massless lines. This is done by the Julia program, using Kruskal's algorithm ([Kruskal 1956](#)) to find the maximal-mass spanning tree of the graph, and then setting the loop momenta to be those internal edges that are missing from the tree. The maximal mass spanning tree of the box is visualised in Figure 5.2 .Using FORM, we can contract indices and apply the momentum conservation rules. Finally, we also apply the relabelling rules described in eq. 5.113. The FORM version of the box integrand is then given in Listing 5.1.

When applied to all 4 diagrams, we obtain precisely the same integrands as those in Kosower, Maybee, and O'Connell ([2019](#)). Additionally, we notice that the two triangles and the box diagram have overlapping propagators. In fact the denominators of the integrands for all three can be written as:

$$\square_{i_1,i_2,i_3,i_4} = \frac{1}{\rho_1^{i_1} \rho_2^{i_2} \rho_3^{i_3} \rho_4^{i_4}}, \quad (5.126)$$

for powers  $\{i\} \in \{0, 1\}^4$ , where the inverse propagators  $\rho_i$  are:

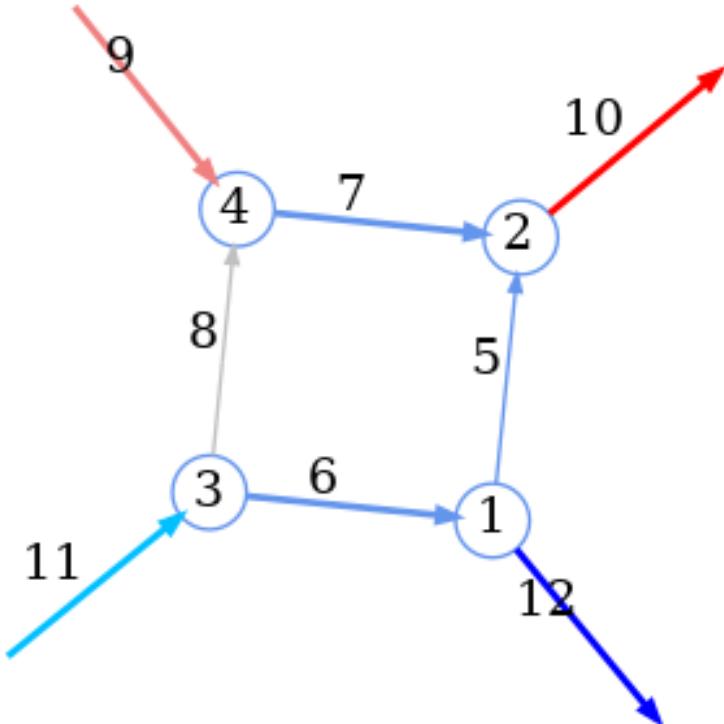


Fig 5.2: The maximal mass spanning tree of the box diagram in blue. The loop momenta are the edges that are missing from the tree.

$$\rho_1 = \ell^2 + i\epsilon, \quad \rho_2 = (\ell - q)^2 + i\epsilon, \quad \rho_3 = \left(\frac{q}{2} - \ell + \tilde{p}_1\right)^2 - m_1^2 + i\epsilon, \quad \rho_4 = \left(-\frac{q}{2} + \ell + \tilde{p}_2\right)^2 - m_2^2 + i\epsilon \quad (5.127)$$

Note that in the classical limit we can expand the massive propagators as shown in eq. 5.109 and obtain, to first order in  $\hbar$ ,

$$\rho_3 = 2\ell \cdot \tilde{p}_1 + i\epsilon, \quad \rho_4 = -2\ell \cdot \tilde{p}_2 + i\epsilon, \quad (5.128)$$

where we used eq. 5.114. The cross box can in fact extend this family, by adding just one more inverse propagator. We now have the complete description of the denominators of all 1-loop contributing diagrams. The numerators in fact do not complicate things much. They are all composed of scalar products of momenta. If these do not contain loop-momenta they are just constant factors in front of the integral. The scalar products that do contain loop momenta can be written as inverse powers of the propagators. This is the case for any Feynman like integrand. Consider the following general integrand:

---

**Listing 5.1** Box integrand in FORM

---

```
[d3l1|o.3.4|i.3.4|i|o|] =
  ( prop(-1,0,q - l1) )
  * ( prop(-1,0,l1) )
  * ( prop(-1,1,1/2*q - l1 + pb1) )
  * ( prop(-1,2, - 1/2*q + l1 + pb2) )
  * ( dot(q,q) - 2*dot(q,l1) - 2*dot(q,pb2) + 2*dot(q,pb1) + dot(l1,l1)
    + 2*dot(l1,pb2) - 2*dot(l1,pb1) - 4*dot(pb2,pb1) )
  * ( dot(l1,l1) + 2*dot(l1,pb2) - 2*dot(l1,pb1) - 4*dot(pb2,pb1) )
  * ( qch2 ) * ( qch2 ) * ( qch1 ) * ( qch1 )
  * ( e ) * ( e ) * ( e ) * ( e );
```

---

$$\mathcal{J} = \frac{\mathcal{N}}{\mathcal{D}} = \frac{\prod_{S'} N_{S'}}{\prod_S D_S} = \frac{\prod_{S'} \sum_{i \geq j} S'_{ij} p_i \cdot p_j}{\prod_S \sum_{i \geq j} S_{ij} p_i \cdot p_j}, \quad (5.129)$$

where  $S$  and  $S'$  are sets of coefficient matrices for all possible dot products. If we have a quadruplet  $\{A, B, a, b\}$  such that  $A_{a,b} \neq 0$  and  $B_{a,b} \neq 0$  and  $p_a$  or  $p_b$  is a loop momentum  $\ell$ , say  $p_a = \ell$ , then it is useful to write this component of the integrand as:

$$\frac{N_B}{D_A} = \frac{\tilde{N}_B + B_{ab}\ell \cdot p_b}{A_{ab}\ell \cdot p_b + \sum_{\substack{i \geq j \\ i,j \neq a,b}} A_{ij}p_i \cdot p_j} = \frac{\tilde{N}_B}{D_A} + \frac{B_{ab}}{A_{ab}} \left( 1 - \frac{\tilde{D}_A}{D_A} \right), \quad (5.130)$$

where the tilde means that we have removed the  $ab$  contribution. Thus, we see that we obtain a sum of objects in the numerator, none of which contains the dot product:  $\ell \cdot p_b$ . We can do this for every quadruplet  $\{A, B, a, b\}$ , after every replacement, and obtain an integrand containing only so called *irreducible* loop-momenta scalar products. In the cases we have here, they all disappear, leaving us truly with a family of integrands like eq. 5.126. This can be readily fed into a IBP solver, such as AIR ([C. Anastasiou and Lazopoulos 2004](#)), FIRE ([Smirnov and Chukharev 2020](#)) or Kira ([Maierhöfer and Usovitsch 2018](#)).

With such a solver we can obtain a complete family of integrals, which can be evaluated. We are thus able to obtain the integrated integrands at NLO that are then used in eq. 5.90 and eq. 5.86. The

algorithm described above was implemented in FORM and is able to rapidly process the 1-loop diagrams. The code is available in the repository for this thesis.

Importantly, we still need to take the classical limit. At this point a problem shows up. The box and cross box integrals are divergent in the classical limit (some authors call this super classical). Let us see this for the box integral. We have from Listing 5.1 that the integral for the box diagram  $B$  is:

$$iB = e^4 Q_1 Q_2 \int d^D \ell p a 4 \tilde{p}_1 \cdot \tilde{p}_2 + 2\ell \cdot (\tilde{p}_2 - \tilde{p}_1) - (q - \ell)^2 (4\tilde{p}_1 \cdot \tilde{p}_2 + 2\ell \cdot (\tilde{p}_2 - \tilde{p}_1) - (\ell)^2) Box_{1,1,1,1} \quad (5.131)$$

we can apply the procedure described above, turning messenger momenta into wave-numbers:

$$iB = e^4 Q_1 Q_2 \int d^D \bar{\ell} \frac{\hbar^D}{\hbar^6} (4\tilde{p}_1 \cdot \tilde{p}_2 + \hbar 2\bar{\ell} \cdot (\tilde{p}_2 - \tilde{p}_1) - \hbar^2 (\tilde{q} - \bar{\ell})^2) (4\tilde{p}_1 \cdot \tilde{p}_2 + \hbar 2\bar{\ell} \cdot (\tilde{p}_2 - \tilde{p}_1) - \hbar^2 (\bar{\ell})^2) \square_{\bar{i}, \bar{i}, \bar{i}, \bar{i}}, \quad (5.132)$$

where  $\square_{\bar{i}_1, \bar{i}_2, \bar{i}_3, \bar{i}_4} = \frac{1}{\hbar^{2(i_1+i_2)+i_3+i_4}} \square_{i_1, i_2, i_3, i_4}$  is the box family with wave-number instead of messenger momenta, and  $\hbar$ s extracted. We notice that if we apply the scalar product reduction eq. 5.130, we will necessarily still have an integral of the form <sup>59</sup>:

$$e^4 Q_1 Q_2 16 (\tilde{p}_1 \cdot \tilde{p}_2)^2 \int d^D \ell \square_{1,1,1,1} = \mathcal{O}(\hbar^{D-6}). \quad (5.133)$$

We need to cancel with the  $\hbar$  in the classical limit (eq. 5.93), here we have one too many orders in the denominator (if  $D=4$ ). Thus, we would get a divergent ( $\frac{1}{\hbar}$ ) contribution when taking the classical  $\hbar \rightarrow 0$  limit. If we consider now the cross box, which is the same family as above but with an extra minus sign:

$$\tilde{p}_4 = \left( \frac{q}{2} - \ell + \tilde{p}_2 \right)^2 - m_2^2 + i\epsilon \simeq 2\ell \cdot \tilde{p}_2 + i\epsilon. \quad (5.134)$$

If we write that  $\square_{i_1^+, i_2^+, i_3^+, i_4^+} = \square_{i_1, i_2, i_3, i_4}$  where the sign is the sign of the  $i\epsilon$  prescription, then we can write the cross box integral as:

$$\int d^D \ell - \square_{1,1,1,1^-} \quad (5.135)$$

<sup>59</sup> along with ones that are of the form  $\square_{1,1,1,0}$  and other permutations. These are not divergent in the classical limit, as they have one less propagator, thus one less  $\hbar$ , just enough to cancel what is left of the  $\hbar$  in the numerator, and have an overall  $\mathcal{O}(\hbar^0)$  scaling.

thus summing up the two divergent contributions from the box and cross box we have:

$$\square_{1,1,1,1} - \square_{1,1,1,1^-} = \square_{1,1,1,0} \left( \frac{1}{-2\ell \cdot \tilde{p}_2 + i\epsilon} - \frac{1}{2\ell \cdot \tilde{p}_2 + i\epsilon} \right) = -i \square_{1,1,1,0} \tilde{\delta}(2\ell \cdot \tilde{p}_2), \quad (5.136)$$

where in the equation line we used *reverse unitarity* (Charalampos Anastasiou and Melnikov 2002; Charalampos Anastasiou, Dixon, and Melnikov 2003; Charalampos Anastasiou et al. 2003). This idea was developed in the context of cross-section calculations in the for collider physics. It enables to set on equal footing real contributions<sup>60</sup> where we integrate over on shell momenta<sup>61</sup> and virtual integrals, where the integration is over all possible four-momentum<sup>62</sup>. The idea is to trade the on-shell delta functions and their  $n$ -th derivatives for differences of (powers of) propagators with alternating  $i\epsilon$ :

$$\frac{i}{(-1)^n n!} \frac{d^n}{dz^n} \tilde{\delta}(z) = \frac{1}{(z - i\epsilon)^{(n+1)}} - \frac{1}{(z + i\epsilon)^{(n+1)}}. \quad (5.137)$$

In our case eq. 5.137 implies the following identities:

$$\tilde{\delta}(-2\ell \cdot \tilde{p}_2) = i(\square_{0,0,0,1} - \square_{0,0,0,1^-}) \quad (5.138)$$

$$\tilde{\delta}(2\ell \cdot \tilde{p}_1) = i(\square_{0,0,1,0} - \square_{0,0,1^-,0}) \quad (5.139)$$

Since we were able to write this difference of propagators as a sort of cut, we could further compare the divergent part of the box and cross box with the corresponding real contribution, which is at tree level, with no cut messengers:  $L' = 0$  and  $|X| = 0$  such that eq. 5.92 is satisfied. We have:

$$\begin{aligned} & \int d^4 \bar{w} \tilde{\delta}(2\tilde{p}_1 \cdot \bar{w} - \hbar \frac{\bar{q}}{2} \cdot \bar{w} + \hbar \bar{w}^2) \tilde{\delta}(-2\tilde{p}_2 \cdot \bar{w} + \hbar \frac{\bar{q}}{2} \cdot \bar{w} + \hbar \bar{w}^2) \\ & \times \hbar \bar{w}^\mu \\ & \times \kappa^4 \hbar^2 \bar{\mathcal{A}}^{(0)}(\tilde{p}_1 - \hbar \frac{\bar{q}}{2}, \tilde{p}_2 + \hbar \frac{\bar{q}}{2} \rightarrow \tilde{p}_1 + \hbar \frac{\bar{q}}{2} + \hbar \bar{w}, \tilde{p}_2 - \hbar \frac{\bar{q}}{2} - \hbar \bar{w}_2) \\ & \times \bar{\mathcal{A}}^{*(0)}(\tilde{p}_1 + \hbar \frac{\bar{q}}{2}, \tilde{p}_2 + \hbar \frac{\bar{q}}{2} - \hbar \bar{q} \rightarrow \tilde{p}_1 - \hbar \frac{\bar{q}}{2} + \hbar \bar{w}, \tilde{p}_2 + \hbar \frac{\bar{q}}{2} - \hbar \bar{w}_2). \end{aligned} \quad (5.140)$$

<sup>60</sup> such as ones in the real integrand eq. 5.51

<sup>61</sup> thus three dimensional momentum space

<sup>62</sup> as virtual particles can be off shell

where the  $\tilde{\mathcal{A}}^{(0)}$  we already computed in eq. 5.118, and we have eliminated the theta functions in preparation of taking the classical limit. Thus, it is essentially a cut of a one-loop box. Notice that in the classical limit we must cancel with  $\hbar^3$  (eq. 5.94) which at leading order is again overdone. The  $\tilde{p}_1 \cdot \tilde{p}_2$  contribution from each tree scales like  $\frac{1}{\hbar^4}$ , and is thus classically divergent as the box and cross box before it. We also notice it has the following form:

$$e^4 Q_1 Q_2 16(\tilde{p}_1 \cdot \tilde{p}_2)^2 \frac{1}{\hbar} \int \tilde{d}^4 \bar{\ell} \bar{\ell}^\mu \tilde{\delta}(2\tilde{p}_1 \cdot \bar{\ell}) \tilde{\delta}(-2\tilde{p}_2 \cdot \bar{\ell}) \square_{\bar{1}, \bar{1}, \bar{0}, \bar{0}}, \quad (5.141)$$

where we have changed variables from  $\bar{w}$  to  $\bar{\ell}$  and removed the  $\hbar(\frac{\bar{q}}{2} \cdot \bar{\ell} \pm \bar{\ell}^2)$  term in the delta functions, in preparation for the classical limit. We can now apply reverse unitarity (eq. 5.138, eq. 5.139) and focus on the integral without the pre-factors:

$$\begin{aligned} \int \tilde{d}^4 \bar{\ell} \bar{\ell}^\mu \tilde{\delta}(2\tilde{p}_1 \cdot \bar{\ell}) \tilde{\delta}(-2\tilde{p}_2 \cdot \bar{\ell}) \square_{\bar{1}, \bar{1}, \bar{0}, \bar{0}} &= \int \tilde{d}^4 \bar{\ell} \bar{\ell}^\mu \tilde{\delta}(2\tilde{p}_1 \cdot \bar{\ell}) i(\square_{0,0,0,1} - \square_{0,0,0,1^-}) \square_{\bar{1}, \bar{1}, \bar{0}, \bar{0}} \\ &= - \int \tilde{d}^4 \bar{\ell} \bar{\ell}^\mu (\square_{0,0,1,0} - \square_{0,0,1^-,0}) (\square_{1,1,0,1} - \square_{1,1,0,1^-}), \\ &= \int \tilde{d}^4 \bar{\ell} \bar{\ell}^\mu [(\square_{1,1,1,1^-} - \square_{1,1,1,1}) + (\square_{1,1,1^-,1} - \square_{1,1,1^-,1^-})]. \end{aligned} \quad (5.142)$$

We now notice something interesting, this cut integral has a sort of horizontal flip symmetry, i.e., if we average over the existing ‘loop’ momentum labelling and a new  $\ell' = q - \ell$  we eliminate the  $\ell^\mu$  dependence in the numerator:

$$\frac{1}{2} [\ell^\mu + (q^\mu - \ell^\mu)] = q^\mu \quad (5.143)$$

since the box families transform as:

$$\square'_{i,j,m,n} = (-1)^m (-1)^n \square_{j,i,-m,-n}. \quad (5.144)$$

Thus we have that the cut integral with the  $\bar{q}^\mu$  factored out:

$$\int \tilde{d}^4 \bar{\ell} \frac{1}{2} [(\square_{1,1,1,1^-} - \square_{1,1,1,1}) + (\square_{1,1,1^-,1} - \square_{1,1,1^-,1^-})]. \quad (5.145)$$

We can apply the relabeling only this time only for the last two integrands, and the integral becomes, applying eq. 5.144:

$$\int \tilde{d}^4\bar{\ell} \frac{1}{2} [(\square_{1,1,1,1^-} - \square_{1,1,1,1}) + (\square'_{1,1,1^-,1} - \square'_{1,1,1^-,1^-})] = \int \tilde{d}^4\bar{\ell} (\square_{1,1,1,1^-} - \square_{1,1,1,1}). \quad (5.146)$$

If we now add the box and cross box contributions from eq. 5.136 we get:

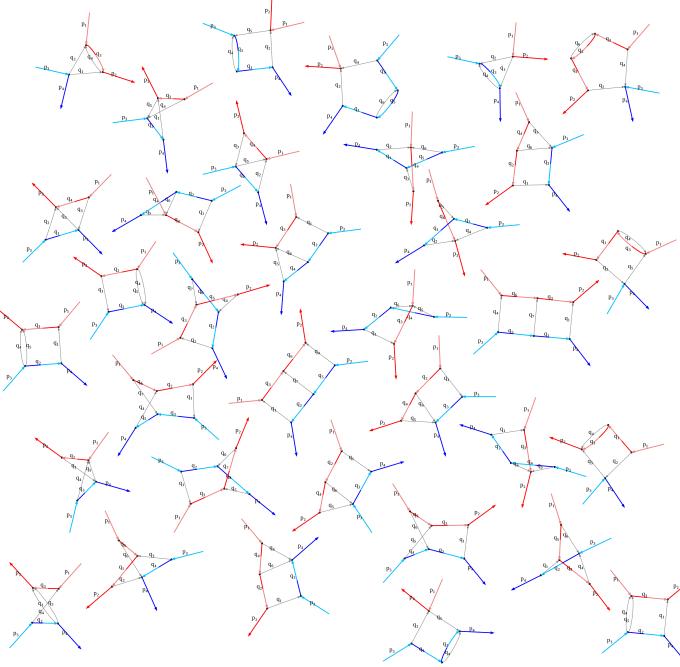
$$\int \tilde{d}^4\bar{\ell} (\square_{1,1,1,1^-} - \square_{1,1,1,1}) + \square_{1,1,1,1} - \square_{1,1,1,1^-} = 0. \quad (5.147)$$

The classically divergent terms have cancelled leaving us with only the finite terms. We now have full control over the classical limit of the one loop contribution to the real and virtual parts of the amplitude.

#### 5.5.4 Higher order and gravity

The code provided is fully general in loop number and can be extended readily to higher loops, at the un-integrated level. The number of contributing diagrams quickly increases however. At two loops, we go from 5 non-scaleless diagrams to 34 (see Figure 5.5.4),

and at three loops we have 470 diagrams (see Figure 5.3).



#### All scaleful diagrams, in the classical limit at two loops in SQED

The classically divergent term cancellations, have been explicitly derived above for the one loop case, however it is not immediately extendable to higher loops. Arguments using Cutkosky rules have been used in (Herrmann et al. [Herrmann et al. 2021](#)) to show this at two loops.

We have up to now only considered the toy model of scalar QED, however the same techniques can be applied to gravity. In this case one considers the same scalars minimally coupled to gravity, using the Einstein-Hilbert action. Going through the same procedures as above, with more algebra, one can also obtain Feynman rules. One key difference is that the gravitons self interact, and do so for any vertex degree. Additionally, the graviton-scalar vertex can also involve any number of gravitons. Of course the highest degree vertex is always limited by the number of loops, such that in practice, one only needs to consider truncated Feynman rules. We implemented these rules in `QGRAF` and `Julia`, and we can see the resulting diagrams at one loop in Figure 5.5.4 only contains one new diagram, and two loops Figure 5.5.4, there are substantially

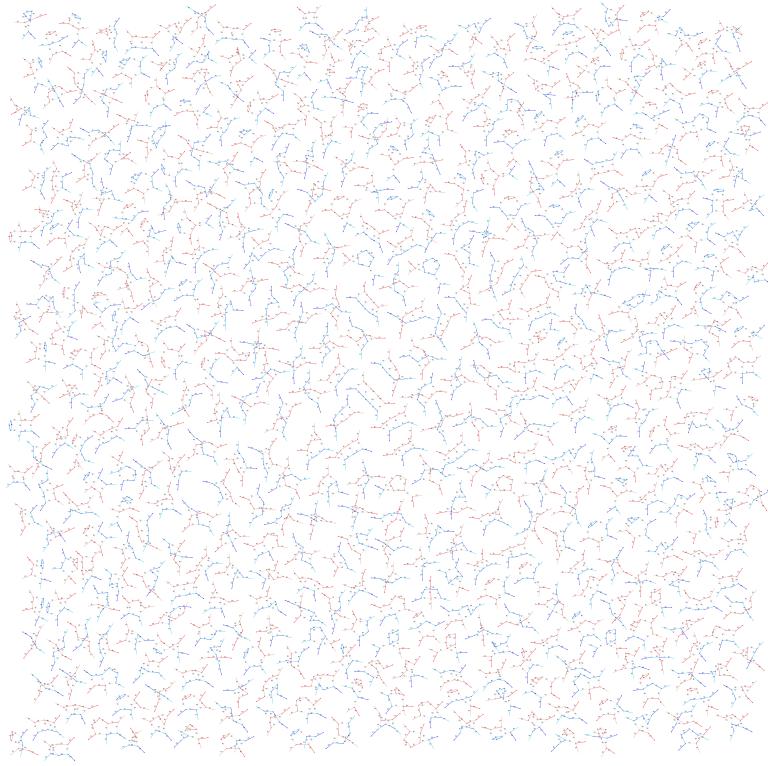
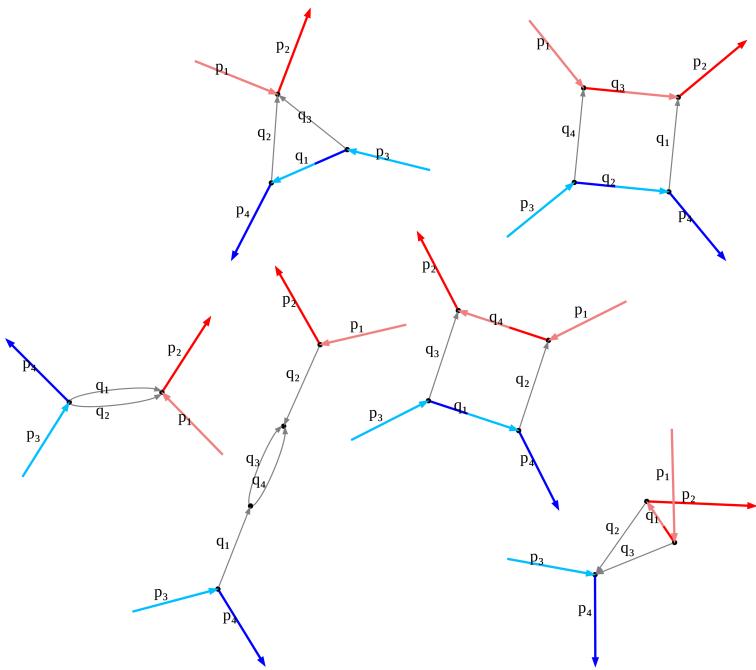
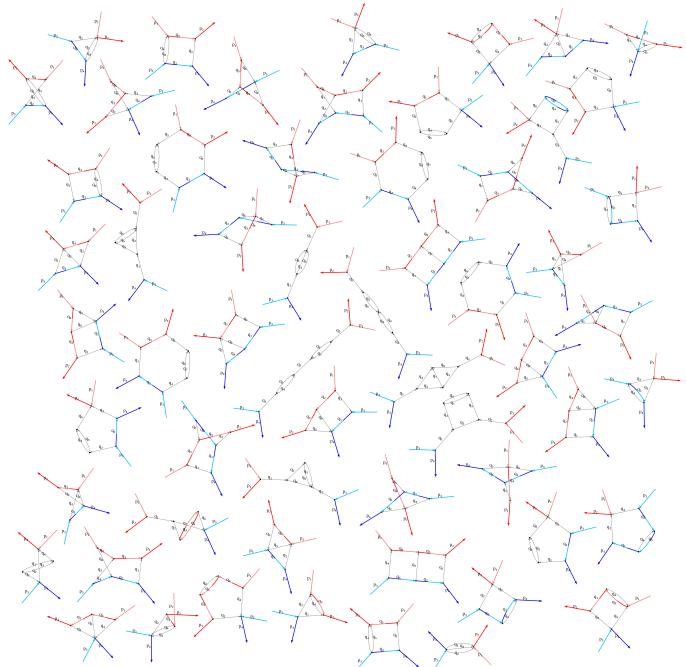


Fig 5.3: All scaleful diagrams, in the classical limit at three loops in SQED

more new diagrams.



All scaleful diagrams, in the classical limit at 1 loop in GR



All scaleful diagrams, in the classical limit at 2 loops in GR

## 6 Conclusion

In this thesis, we explored the landscape of theoretical research that surrounds gravitational wave detection. We showed the existence of gravitational waves theoretically. We motivated waveform generation techniques grounded in first principles' physics. It is only through such precise frameworks that we can hope to elucidate the nature of the objects that emit gravitational radiation. We looked at NR and EOB two complementary approaches to waveform generation that enabled LIGO to make the field changing detection in 2016. Both approaches have their challenges, NR with stability and computational cost, and EOB with higher order perturbative calculations. Each increase in precision enables a more precise match of the signal, enabling a much richer analysis of the incoming wave. Purely analytical tools have recently come into the crosshairs of particle physicists as well.

We then went on to understand these nascent formalisms built on techniques from particle physics. The EFT matching approach enables a simple map from EOB to amplitudes computed in the potential kinematic region. EFT, and EOB in general, struggle to use and implement dissipation however. The Hamiltonian framework has to be somehow augmented with the dissipative forces. This motivates frameworks that sidestep the EOB framework. We explored the most developed one currently: KMOC . We were able to express classical observables such as the impulse in terms of the classical limit amplitudes and their unitarity cuts. We then went on to implement a programmatic framework for expressing the necessary integrals that the amplitudes yield. We used custom code to apply Feynman rules, manipulate and filter the graphs. We matched the results from Kosower, Maybee, and O'Connell (2019), and then implemented a scalar product reduction for preparing the input to IBP reduction programs.

The results and formalisms explored here show that amplitude techniques in gravitational wave theory are a very promising and powerful tool. Much of the recent HEP techniques can directly

be put to use in this context, and can already outperform traditional methods. The higher perturbative order frontier however is increasingly difficult. At some point in the near future, these computations will have exhausted the techniques currently in use for particle physics. The next step is to develop new techniques that can be applied to both fields. Important work needs to be done to have a fully amplitude based waveform-generation framework. This might be the next revolution in gravitational wave theory.

## References

- Abbott, B. P., R. Abbott, T. D. Abbott, et al. 2016. “Observation of Gravitational Waves from a Binary Black Hole Merger.” *Physical Review Letters* 116 (6): 061102. <https://doi.org/10.1103/PhysRevLett.116.061102>.
- Abbott, B. P., Rich Abbott, Thomas D. Abbott, et al. 2017. “GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral.” *Physical Review Letters* 119 (16): 161101. <https://doi.org/10.1103/PhysRevLett.119.161101>.
- Abbott, R., T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, et al. 2021. “Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences.” *The Astrophysical Journal Letters* 915 (1): L5. <https://doi.org/10.3847/2041-8213/ac082e>.
- Anastasiou, Charalampos, Ruth Britto, Bo Feng, Zoltan Kunszt, and Pierpaolo Mastrolia. 2007. “D-Dimensional Unitarity Cut Method.” *Physics Letters B* 645 (2-3): 213–16. <https://doi.org/10.1016/j.physletb.2006.12.022>.
- Anastasiou, Charalampos, Lance Dixon, and Kirill Melnikov. 2003. “NLO Higgs Boson Rapidity Distributions at Hadron Colliders.” *Nuclear Physics B - Proceedings Supplements*, Proceedings of the 6th International Symposium on Radiative Corrections and the 6th Zeuthen Workshop on Elementary Particle Theory, 116: 193–97. [https://doi.org/10.1016/S0920-5632\(03\)80168-8](https://doi.org/10.1016/S0920-5632(03)80168-8).
- Anastasiou, Charalampos, Lance Dixon, Kirill Melnikov, and Frank Petriello. 2003. “Dilepton Rapidity Distribution in the Drell-Yan Process at Next-to-Next-to-Leading Order in QCD.” *Physical Review Letters* 91 (18): 182002. <https://doi.org/10.1103/PhysRevLett.91.182002>.
- Anastasiou, Charalampos, and Kirill Melnikov. 2002. “Higgs Boson Production at Hadron Colliders in NNLO QCD.” *Nuclear Physics B* 646 (1): 220–56. [https://doi.org/10.1016/S0550-3213\(02\)00837-4](https://doi.org/10.1016/S0550-3213(02)00837-4).
- Anastasiou, C., and A. Lazopoulos. 2004. “Automatic Integral Reduction for Higher Order Perturbative Calculations.” *Journal*

- of High Energy Physics* 2004 (07): 046–46. <https://doi.org/10.1088/1126-6708/2004/07/046>.
- Bern, Z., C. Cheung, R. Roiban, Chia-Hsien Shen, M. P. Solon, and M. Zeng. 2019. “Black Hole Binary Dynamics from the Double Copy and Effective Theory.” *Journal of High Energy Physics* 10 (October): 206. [https://doi.org/10.1007/JHEP10\(2019\)206](https://doi.org/10.1007/JHEP10(2019)206).
- Bern, Z., L. Dixon, D. C. Dunbar, and D. A. Kosower. 1995. “Fusing Gauge Theory Tree Amplitudes Into Loop Amplitudes.” *Nuclear Physics B* 435 (1-2): 59–101. [https://doi.org/10.1016/0550-3213\(94\)00488-Z](https://doi.org/10.1016/0550-3213(94)00488-Z).
- Bern, Zvi, Vittorio Del Duca, Lance J. Dixon, and David A. Kosower. 2005. “All Non-Maximally-Helicity-Violating One-Loop Seven-Gluon Amplitudes in N=4 Super-Yang-Mills Theory.” *Physical Review D* 71 (4): 045006. <https://doi.org/10.1103/PhysRevD.71.045006>.
- Bern, Zvi, Lance Dixon, and David A. Kosower. 1994. “Dimensionally-Regulated Pentagon Integrals.” *Nuclear Physics B* 412 (3): 751–816. [https://doi.org/10.1016/0550-3213\(94\)90398-0](https://doi.org/10.1016/0550-3213(94)90398-0).
- Bern, Zvi, Juan Pablo Gatica, Enrico Herrmann, Andres Luna, and Mao Zeng. 2022. “Scalar QED as a Toy Model for Higher-Order Effects in Classical Gravitational Scattering.” *arXiv:2112.12243 [Gr-Qc, Physics:hep-Th]* 08 (August): 131. [https://doi.org/10.1007/JHEP08\(2022\)131](https://doi.org/10.1007/JHEP08(2022)131).
- Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B. Shah. 2015. “Julia: A Fresh Approach to Numerical Computing.” *arXiv*. <https://arxiv.org/abs/1411.1607>.
- Birkhoff, George David, and Rudolph Ernest Langer. 1923. *Relativity and Modern Physics*. Cambridge: Harvard University Press; [etc., etc.].
- Bjerrum-Bohr, N. E. J., John F. Donoghue, and Pierre Vanhove. 2014. “On-Shell Techniques and Universal Results in Quantum Gravity.” *Journal of High Energy Physics* 02 (2): 111. [https://doi.org/10.1007/JHEP02\(2014\)111](https://doi.org/10.1007/JHEP02(2014)111).
- Blanchet, L., and T. Damour. 1986. “Radiative Gravitational Fields in General Relativity I. General Structure of the Field Outside the Source.” *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 320 (1555): 379–430.
- Blanchet, Luc. 2014. “Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries.” *Living Reviews in Relativity* 17 (1): 2. <https://doi.org/10.12942/lrr-2014-2>.

- Blanchet, Luc, Thibault Damour, and Bala R. Iyer. 1995. “Gravitational Waves from Inspiralling Compact Binaries: Energy Loss and Waveform to Second-Post-Newtonian Order.” *Physical Review D* 51 (10): 5360. <https://doi.org/10.1103/PhysRevD.51.5360>.
- Bonnor, W. B., and Louis Rosenhead. 1959. “Spherical Gravitational Waves.” *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 251 (994): 233–71. <https://doi.org/10.1098/rsta.1959.0003>.
- Brezin, E., C. Itzykson, and J. Zinn-Justin. 1970. “Relativistic Balmer Formula Including Recoil Effects.” *Physical Review D* 1 (8): 2349–55. <https://doi.org/10.1103/PhysRevD.1.2349>.
- Buchbinder, Evgeny I., and Freddy Cachazo. 2005. “Two-Loop Amplitudes of Gluons and Octa-Cuts in N=4 Super Yang-Mills.” *Journal of High Energy Physics* 11 (11): 036. <https://doi.org/10.1088/1126-6708/2005/11/036>.
- Buonanno, A., and T. Damour. 1999. “Effective One-Body Approach to General Relativistic Two-Body Dynamics.” *Physical Review D* 59 (8): 084006. <https://doi.org/10.1103/PhysRevD.59.084006>.
- Buonanno, Alessandra, and Thibault Damour. 2000. “Transition from Inspiral to Plunge in Binary Black Hole Coalescences.” *Physical Review D* 62 (6): 064015. <https://doi.org/10.1103/PhysRevD.62.064015>.
- Cachazo, Freddy, and Alfredo Guevara. 2020. “Leading Singularities and Classical Gravitational Scattering” 02 (February): 181. [https://doi.org/10.1007/JHEP02\(2020\)181](https://doi.org/10.1007/JHEP02(2020)181).
- Campanelli, M., C. O. Lousto, P. Marronetti, and Y. Zlochower. 2006. “Accurate Evolutions of Orbiting Black-Hole Binaries Without Excision.” *Physical Review Letters* 96 (11): 111101. <https://doi.org/10.1103/PhysRevLett.96.111101>.
- Carroll, Sean M. 2019. “Spacetime and Geometry: An Introduction to General Relativity.” *Higher Education from Cambridge University Press*. <https://www.cambridge.org/highereducation/books/spacetime-and-geometry/38ED-ABF9E2BADCE6FBCF2B22DC12BFFE>; Cambridge University Press. <https://doi.org/10.1017/9781108770385>.
- Cheung, Clifford, Ira Z. Rothstein, and Mikhail P. Solon. 2018. “From Scattering Amplitudes to Classical Potentials in the Post-Minkowskian Expansion.” *Physical Review Letters* 121 (25): 251101. <https://doi.org/10.1103/PhysRevLett.121.251101>.
- Coleman, Sidney. 2018. *Lectures of Sidney Coleman on Quantum Field Theory*. Edited by Bryan Gin-ge Chen, David Derbes,

- David Griffiths, Brian Hill, Richard Sohn, and Yuan-Sen Ting. Hackensack: WSP. <https://doi.org/10.1142/9371>.
- Collins, John. 2019. “A New Approach to the LSZ Reduction Formula,” April. <https://arxiv.org/abs/1904.10923>.
- Cristofoli, Andrea, Riccardo Gonzo, David A. Kosower, and Donal O’Connell. 2022. “Waveforms from Amplitudes.” *Phys. Rev. D* 106 (5): 056007. <https://doi.org/10.1103/PhysRevD.106.056007>.
- Damour, Thibault. 2016. “Gravitational Scattering, Post-Minkowskian Approximation and Effective One-Body Theory.” *Physical Review D* 94 (10): 104015. <https://doi.org/10.1103/PhysRevD.94.104015>.
- Detweiler, S. 1979. “Pulsar Timing Measurements and the Search for Gravitational Waves.” *The Astrophysical Journal* 234: 1100–1104. <https://doi.org/10.1086/157593>.
- Goldberger, Walter D., and Ira Z. Rothstein. 2006. “Effective Field Theory of Gravity for Extended Objects.” *Physical Review D* 73 (10): 104029. <https://doi.org/10.1103/PhysRevD.73.104029>.
- Goldstein, Herbert, Charles P. Poole, and John L. Safko. 2002. *Classical Mechanics*. Addison Wesley.
- Guevara, Alfredo. 2019. “Holomorphic Classical Limit for Spin Effects in Gravitational and Electromagnetic Scattering.” *Journal of High Energy Physics* 04 (4): 033. [https://doi.org/10.1007/JHEP04\(2019\)033](https://doi.org/10.1007/JHEP04(2019)033).
- Hahn, Susan G., and Richard W. Lindquist. 1964. “The Two-Body Problem in Geometrodynamics.” *Annals of Physics* 29 (September): 304–31. [https://doi.org/10.1016/0003-4916\(64\)90223-4](https://doi.org/10.1016/0003-4916(64)90223-4).
- Herrmann, Enrico, Julio Parra-Martinez, Michael S. Ruf, and Mao Zeng. 2021. “Radiative Classical Gravitational Observables at  $\mathcal{O}(G^3)$  from Scattering Amplitudes.” *Journal of High Energy Physics* 10 (10): 148. [https://doi.org/10.1007/JHEP10\(2021\)148](https://doi.org/10.1007/JHEP10(2021)148).
- Hida, Kichiro, and Hiroshi Okamura. 1972. “Gauge Transformation and Gravitational Potentials.” *Progress of Theoretical Physics* 47 (5): 1743–57. <https://doi.org/10.1143/PTP.47.1743>.
- Hobbs, G., A. Archibald, Z. Arzoumanian, D. Backer, M. Bailes, N. D. R. Bhat, M. Burgay, et al. 2010. “The International Pulsar Timing Array Project: Using Pulsars as a Gravitational Wave Detector.” *Classical and Quantum Gravity* 27 (8): 084013. <https://doi.org/10.1088/0264-9381/27/8/084013>.
- Iwasaki, Yoichi. 1971. “Quantum Theory of Gravitation Vs. Classical Theory\*): Fourth-Order Potential.” *Progress of Theoretical Physics* 46 (5): 1587–1609. <https://doi.org/10.1143/PTP.46.158>

[7.](#)

- Kalin, Gregor, and Rafael A. Porto. 2020a. “From Boundary Data to Bound States.” *Journal of High Energy Physics* 01 (January): 072. [https://doi.org/10.1007/JHEP01\(2020\)072](https://doi.org/10.1007/JHEP01(2020)072).
- . 2020b. “From Boundary Data to Bound States. Part II. Scattering Angle to Dynamical Invariants (with Twist).” *Journal of High Energy Physics* 02 (February): 120. [https://doi.org/10.1007/JHEP02\(2020\)120](https://doi.org/10.1007/JHEP02(2020)120).
- Kälin, Gregor, Jakob Neef, and Rafael A. Porto. 2022. “Radiation-Reaction in the Effective Field Theory Approach to Post-Minkowskian Dynamics,” July. <https://arxiv.org/abs/2207.00580>.
- Kälin, Gregor, and Rafael A. Porto. 2020. “Post-Minkowskian Effective Field Theory for Conservative Binary Dynamics.” *Journal of High Energy Physics* 11 (11): 106. [https://doi.org/10.1007/JHEP11\(2020\)106](https://doi.org/10.1007/JHEP11(2020)106).
- Kosower, David A., Ben Maybee, and Donal O’Connell. 2019. “Amplitudes, Observables, and Classical Scattering.” *Journal of High Energy Physics* 02 (2): 137. [https://doi.org/10.1007/JHEP02\(2019\)137](https://doi.org/10.1007/JHEP02(2019)137).
- Kotikov, A. V. 1991. “Differential Equations Method. New Technique for Massive Feynman Diagram Calculation.” *Physics Letters B* 254 (1): 158–64. [https://doi.org/10.1016/0370-2693\(91\)90413-K](https://doi.org/10.1016/0370-2693(91)90413-K).
- Kruskal, Joseph B. 1956. “On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem.” *Proceedings of the American Mathematical Society* 7 (1): 48–50. <https://doi.org/10.1090/S0002-9939-1956-0078686-7>.
- Kuipers, J., T. Ueda, J. A. M. Vermaasen, and J. Vollinga. 2013. “FORM Version 4.0.” *Comput. Phys. Commun.* 184: 1453–67. <https://doi.org/10.1016/j.cpc.2012.12.028>.
- Lehmann, H., K. Symanzik, and W. Zimmermann. 1955. “Zur Formulierung quantisierter Feldtheorien.” *Il Nuovo Cimento (1955-1965)* 1 (1): 205–25. <https://doi.org/10.1007/BF02731765>.
- Maierhöfer, Philipp, and Johann Usovitsch. 2018. “Kira 1.2 Release Notes.” *arXiv:1812.01491 [Hep-Ph]*, December. <https://arxiv.org/abs/1812.01491>.
- Maiorano, Michele, Francesco De Paolis, and Achille A. Nucita. 2021. “Principles of Gravitational-Wave Detection with Pulsar Timing Arrays.” *Symmetry* 13 (12): 2418. <https://doi.org/10.3390/sym13122418>.
- Mukherjee, Sajal, Sanjit Mitra, and Sourav Chatterjee. 2021. “Grav-

- itational Wave Observatories May Be Able to Detect Hyperbolic Encounters of Black Holes.” *Monthly Notices of the Royal Astronomical Society* 508 (4): 5064–73. <https://doi.org/10.1093/mnras/stab2721>.
- Neill, Duff, and Ira Z. Rothstein. 2013. “Classical Space-times from the S-matrix.” *Nuclear Physics B* 877 (2): 177–89. <https://doi.org/10.1016/j.nuclphysb.2013.09.007>.
- Nogueira, P. 1993. “Automatic Feynman Graph Generation.” *Journal of Computational Physics* 105 (2): 279–89. <https://doi.org/10.1006/jcph.1993.1074>.
- Paton, Keith. 1969. “An Algorithm for Finding a Fundamental Set of Cycles of a Graph.” *Communications of the ACM* 12 (9): 514–18. <https://doi.org/10.1145/363219.363232>.
- Porto, Rafael A. 2016. “The Effective Field Theorist’s Approach to Gravitational Dynamics.” *Physics Reports* 633 (May): 1–104. <https://doi.org/10.1016/j.physrep.2016.04.003>.
- Pretorius, Frans. 2005. “Evolution of Binary Black-Hole Space-times.” *Physical Review Letters* 95 (12): 121101. <https://doi.org/10.1103/PhysRevLett.95.121101>.
- Remiddi, E. 1997. “Differential Equations for Feynman Graph Amplitudes.” *Il Nuovo Cimento A (1971-1996)* 110 (12): 1435–52. <https://doi.org/10.1007/BF03185566>.
- Shomer, Assaf. 2007. “A Pedagogical Explanation for the Non-Renormalizability of Gravity,” December. <https://arxiv.org/abs/0709.3555>.
- Smirnov, A. V., and F. S. Chukharev. 2020. “FIRE6: Feynman Integral REduction with Modular Arithmetic.” *Computer Physics Communications* 247 (February): 106877. <https://doi.org/10.1016/j.cpc.2019.106877>.
- Srednicki, Mark Allen. 2007. *Quantum Field Theory*. Cambridge ; New York: Cambridge University Press.
- Sturani, Riccardo. 2021. “Effective Field Theory Methods to Model Compact Binaries.” In *Handbook of Gravitational Wave Astronomy*, edited by Cosimo Bambi, Stavros Katsanevas, and Konstantinos D. Kokkotas, 1–33. Singapore: Springer. [https://doi.org/10.1007/978-981-15-4702-7\\_32-1](https://doi.org/10.1007/978-981-15-4702-7_32-1).
- ’t Hooft, G., and M. Veltman. 1972. “Regularization and Renormalization of Gauge Fields.” *Nuclear Physics B* 44 (1): 189–213. [https://doi.org/10.1016/0550-3213\(72\)90279-9](https://doi.org/10.1016/0550-3213(72)90279-9).
- Thorne, Kip S. 1980. “Multipole Expansions of Gravitational Radiation.” *Reviews of Modern Physics* 52 (2): 299–339. <https://doi.org/10.1103/RevModPhys.52.299>.
- Vaidya, Varun. 2015. “Gravitational Spin Hamiltonians from the S

- Matrix." *Physical Review D* 91 (2): 024017. <https://doi.org/10.1103/PhysRevD.91.024017>.
- Vermaseren, J. A. M. 2000. "New Features of FORM," October. <https://arxiv.org/abs/math-ph/0010025>.
- Weinberg, Steven. 1972. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: Wiley.