Gravitational Waves
Interplay Between Mathematical Foundations & Observations

Abhay Ashtekar
Institute for Gravitation & the Cosmos and Physics Department
The Pennsylvania State University

Abstract

Since the first detection of Gravitational Waves ~8 years ago, the field has literally exploded in multiple directions that include multi-wavelength astronomy and astrophysics, approximation methods in general relativity, numerical relativity, applications of machine learning to waveform model building, forefront cosmological issues such as the Hubble tension, and nuclear physics issues related to the equation of state of neutron stars and nuclear processes at extreme temperature. Therefore Gravitational Wave Science has emerged as one of the most exciting fields to work on, that now attracts young researchers in large numbers.

At the same time the very explosion of the field makes it difficult for these researchers to grasp, *even in broad terms,* the conceptual and mathematical foundation of the theory of Gravitational Waves since these foundations are rarely discussed in the specific areas these researchers work in everyday. The purpose of my lectures is to fill this gap.

As I will discuss, the notion of \`radiation' requires global and rather subtle constructions. For several decades there was considerable confusion even about the physical reality of gravitational waves in full general relativity! This confusion was dispelled, thanks to a beautiful interplay between physics and geometry. I believe that every theoretical researcher in the field should be aware, *at least in general terms*, of the way that difficulties associated with coordinate invariance are overcome and fully gauge invariant mathematical quantities representing physical observables are extracted. This awareness would provide a broad perspective that can guide their own research. Furthermore, as I discuss in the last two lectures, foundational issues can also have concrete applications in addressing \`practical issues'.

Content: The first lecture discusses some subtleties associated with the notion of radiation already for Maxwell fields in Minkowski space. The second lecture shows that the techniques used to define an unambiguous notion of electromagnetic radiation can be directly generalized to gravitational waves in exact general relativity. The third lecture introduces the BMS group as the asymptotic symmetry group and in the fourth lecture I show how these symmetries lead to an infinite set of observables and balance laws they satisfy. In the last two lectures I discuss applications of these balance laws to improve the waveform models, first explaining the need for improvement, and then providing concrete illustrations of these improvements.

PLAN OF THE MINI-COURSE

My lectures will focus on

(i) Part I: Conceptual and Mathematical issues associated with gravitational waves (GWs) in full, non-linear general relativity. They will thus complement other lectures on approximation methods and numerical relativity by providing the concepts and mathematical notions they use;

and,

(ii) Part II: How these results in exact general relativity can be used as diagnostic tools to test the accuracy of model waveforms. Normally one uses numerical simulations to evaluate the accuracy but there are regions of parameter space where numerical simulations are sparse. The diagnostic tests come from identities that must be satisfied in exact GR. Their strength lies in the fact that they enable one to test accuracy of waveform models even when one does not have the exact waveform to compare them with! They focus on the accuracy of the waveform vis a vis (infinitely many) physical observables, thereby bringing out the physical nature of the inaccuracy and suggesting directions for improvements in all regions of the parameter space. Furthermore, they can be used to test accuracy of NR waveforms themselves.

Main References where further details for the material covered can be found:

Lecture 1: Sections 1 and 2 of AA & Bonga, Gen. Rel. Gravit. Grab. 49, 122 (43pp) (2017); https://arxiv.org/pdf/1707.09914.
Sections I and II of Newman and Penrose, Proc. R. Soc (London) 305, 175-204 (1968)

Lecture 2: Sections I and II of AA, in the GR Centennial volume, edited by Beiri & Yau; https://arxiv.org/pdf/1409.1800
Section II (parts 8-11) of R. Penrose, Proc. R. Soc (London) 284, 159-203 (1965)
AA, J. Math. Phys. 22, 2885-2895 (1981)

Lecture 3: http://igpg.gravity.psu.edu/research/asymquant-book.pdf Pages 44-54
AA, De Lorenzo & Khera, Phys. Rev. D101, 044005 (1-17) (2020) https://arxiv.org/pdf/1910.02907.pdf

Lecture 4: http://igpg.gravity.psu.edu/research/asymquant-book.pdf Pages 55-77
Phase space of radiative modes: AA & A. Magnon, Comm. Math. Phys. 86, 55-68 (1982).
BMS. Hamiltonians/fluxes: AA & Streubel, Proc. R. Soc (London) A376, 585-607 (1981)

Lectures 5&6: AA, De Lorenzo & Khera, GRG 52, 107 (1-27) (2020); https://arxiv.org/pdf/1906.00913.pdf
Khera, Krishnan, AA & Del Lorenzo, Phys. Rev. D 103, 044012 (2021); https://arxiv.org/pdf/2009.06351.pdf
Mitman et al, Phys. Rev. D 103, 024031 (2021); https://arxiv.org/pdf/2011.01309.pdf
Khera, AA & Krishnan, https://arxiv.org/pdf/2107.09536.pdf

Students: Doing the exercises is important!

Lecture #1 : Electromagnetic Waves

- "Radiation content" in a solution to Einstein's eqns is tricky to identify

  Led to a lot of confusion about reality of GWs in full GR in the early days. We will see that Einstein himself contributed to this confusion! Clarified in the 1960s and 1970s by Bondi-Sachs, Newman-Penrose and others.

- Some aspects of the confusion are present already for Maxwell fields in Minkowski space. So, in this lecture we will begin with this simpler case. Provides intuition & techniques for GWs we will then study.

- $(M = \mathbb{R}^4, \eta_{ab})$: Minkowski space-time. Maxwell field F_{ab}

  $$\nabla_{[a} F_{bc]} = 0 \quad \& \quad \nabla_a F^{ab} = -4\pi J_b$$

  Non-zero only in a spatially compact world tube

  Main issue : cannot extract the "radiative content" of F_{ab} at a finite distance from sources. No local criterion.

- Poynting vector non-zero : $\vec{E} \times \vec{B} \neq 0$ seems like a natural criterion. But it is not! condition is not Lorentz invariant.

  Ex: coulomb field of a point charge: In the rest frame
  $E^a = \frac{e}{r^2} \hat{r}^a, \quad B^a = 0 \Rightarrow \vec{E} \times \vec{B} = 0.$

  But if you boost e.g. in the z-direction, in the new rest frame
  $\vec{E}' \times \vec{B}' \neq 0$

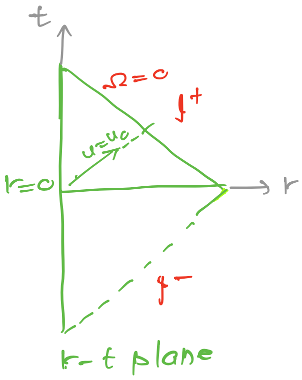
- What is radiation? "$\frac{1}{r}$ part" of the field. But one cannot extract it if you are given the solution only locally.

- Need to go in the "far field" region; mathematically $r \to \infty$.

  "Radiative part" is a global concept.

Further details for material covered lecture #1:
Sections 1 and 2 of AA & Bonga, Gen. Rel. Gravit. 49, 122 (43pp) (2017); https://arxiv.org/pdf/1707.09914.
Sections 1 and 2 of Newman and Penrose, Proc. R. Soc (London) 305, 175-204 (1968)

• A precise way to do this to bring ∞ to a fine distance by an appropriate conformal transformation : Penrose completion.



Mink space : $u=t-r \qquad v=t+r$;

$$ds^2 = -dt^2+dr^2+r^2\underbrace{(d\theta^2+\sin^2\theta\, d\varphi^2)}_{d\omega^2}$$

$$= -du^2 - 2\,du\,dr + r^2 d\omega^2$$

At $r\to\infty$, $u=u_0$, Metric is ill-defined. But

$$d\hat{s}^2 = \Omega^2 ds^2 = \Omega^2 du^2 + du\,d\Omega + d\omega^2 ; \qquad \Omega=\frac{1}{r}$$

is a well-defined metric in a nbd of $\Omega=0$.

$\Omega=0$ not part of Mink space ($r=\infty$ there).

Have attached a boundary to Minkowski space $\Rightarrow$ completion

Boundary called $\mathcal{J}^+$ (scri-plus) : end-points of null geodesics $u=u_0$ in Minkowski space : Natural Home for radiation fields.

$$\text{At } \mathcal{J}^+ : \quad \Omega \hat{=} 0 \Rightarrow d\hat{s}^2 \hat{=} du\,d\Omega + \underbrace{d\omega^2}_{\text{unit 2-sphere metric}}$$

Tgt vectors to $\mathcal{J}^+$: $\frac{\partial}{\partial u}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \varphi} \Rightarrow \frac{\partial}{\partial u}$: Null vector

& although $\mathcal{J}^+$ is 3-dimensional, Intrinsic metric $d\hat{s}^2 = d\omega^2$

signature 0,++ : degenerate

$\mathcal{J}^+$ is null surface, with $\frac{\partial}{\partial u}$ as its null normal.

• We can use advanced null coordinate $v=t+r$ in place of $u=t-r$, then we get past null infinity $\mathcal{J}^-$. This is where we specify "No incoming radiation" condition (Retarded fields).

• Null tetrad : convenient basis to expand fields (Newman-Penrose)

$$n_a = \nabla_a(t-r) \qquad l_a = \tfrac{1}{2}\nabla_a(t+r) \qquad m_a = \frac{r}{\sqrt{2}}(\nabla_a\theta + i\sin\theta\, \nabla_a\varphi)$$

Then: $n_a, l_a, m_a, \bar{m}_a$ are all null, $\begin{cases} n_a l^a = -1 \quad m^a \bar{m}_a = 1 \\ \text{All other contractions vanish.}\end{cases}$

For the rescaled metric $\Omega^2 \eta_{ab} = \hat{\eta}^{ab}$, the null tetrad is

$$\hat{n}^a = n^a \qquad \hat{l}^a = r^2 l^a \equiv \Omega^{-2} l^a , \qquad \hat{m}^a = \Omega^{-1} m^a , \quad \hat{\bar{m}}^a = \Omega^{-1} \bar{m}^a$$

The hatted vector fields have smooth limit to $\mathcal{J}^+$.

unhatted fields : Indices raised & lowered using η^{ab} & η_{ab} and for hatted fields using $\hat{\eta}^{ab}$ & $\hat{\eta}_{ab}$.

• conformal invariance of Maxwell eqns :

$\hat{F}_{ab} = F_{ab}$ satisfies $\hat{\nabla}_{[a}\hat{F}_{bc]} = 0$ & $\hat{\nabla}^a \hat{F}_{bc} = -4\pi \hat{J}_b \qquad (\hat{J}_b = \Omega^{-2} J_b)$

$\hat{J}^b$: C^∞ & spatially compact support

$\mathscr{I}^+$ is a regular sub-manifold in the completed spacetime

$\Rightarrow$ $\hat{F}_{ab} = F_{ab}$ smooth tensor field @ $\mathscr{I}^+$. Hence for components:

$$\Phi_2 := F_{ab}\, n^a \bar{m}^b = \hat{F}_{ab}\, \hat{n}^a \frac{\hat{\bar{m}}^b}{r} = \frac{\Phi_2^\circ}{r} + O\left(\frac{1}{r^2}\right); \qquad \Phi_2^\circ \,\hat{=}\, \hat{F}_{ab}\hat{n}^a\hat{\bar{m}}^b$$

$$\Phi_1 = \tfrac{1}{2}F_{ab}\,(\underset{\hat{n}^a}{n^a}\,\underset{\hat{\ell}^b/r^2}{\ell^b} + \underset{\hat{m}^a/r}{m^a}\,\bar{m}^b) = \frac{\Phi_1^\circ}{r^2} + O\left(\frac{1}{r^3}\right); \qquad \text{etc}$$

$$\Phi_0 = F_{ab}\, m^a \ell^b = \frac{\Phi_0^\circ}{r^3} + O\left(\frac{1}{r^4}\right)$$

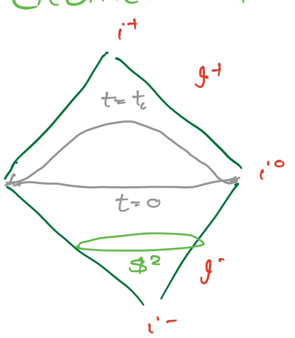
NP components of the Maxwell field — specific fall-off as one $\to \mathscr{I}^+$. "Peeling"

Ex: check that Peeling property holds. If $A_a = \hat{A}_a$ is a potential for $F_{ab} = \hat{F}_{ab}$ in the gauge $\hat{A}_a\hat{n}^a \,\hat{=}\, 0$, show that $(\mathcal{L}_n A_a \hat{\bar{m}}^a) \,\hat{=}\, \Phi_2^\circ$. $A_a m^a$ or Φ_2° are the 2 radiative modes of EM waves.

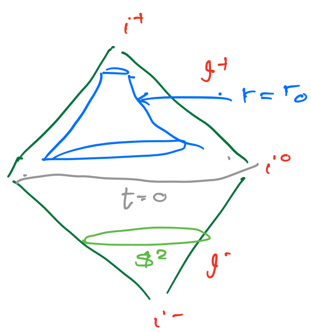
- We return to the question we began with: Radiation content of Maxwell field $\sim$ '$\frac{1}{r}$ part': isolated @ $\mathscr{I}^+$: Φ_2°
- The $\frac{1}{r^2}$ part is the Coulombic part: isolated @ $\mathscr{I}^+$: Φ_1°

- Energy, momentum, Angular momentum carried by EM waves: All expressible as integrals over $\mathscr{I}^+$

Geometrical considerations:



$t = t_0$: spacelike planes $\to \mathscr{I}^+$ as $t\to\infty$



$r = r_0$: Timelike cylinders $\to \mathscr{I}^+$

i^+, i° : $u \to \pm\infty$ ends of $\mathscr{I}^+$; i°, i^- : $v \to \pm\infty$ ends of $\mathscr{I}^-$

Flux across $t = t_0$ planes (or $r = r_0$ cylinders) associated with a Killing vector field K^a:

$$\mathcal{F}_K = \int_{t=t_0} T_{ab}\, K^a\, \underbrace{dS^b}_{(t^a \, d^3x)}$$

$$T_{ab} = F_{am}F_{bn}\, g^{mn} - \text{Trace}$$

$$= \hat{F}_{am}\hat{F}_{bn}\,\Omega^2 \hat{g}^{mn} - \text{Trace}$$

Ex: calculate for K^a = Translations & show:

$$\underbrace{\mathcal{F}_{\vec{t}}}_{\text{energy flux}} = \int_{\mathscr{I}^+} |\Phi_2^\circ|^2\, du\, \underbrace{d^2S}_{\sin\theta\, d\theta\, d\varphi}$$

$$\underbrace{\mathcal{F}_{\vec{z}}}_{\text{Momentum flux}} = \int_{\mathscr{I}^+} \underbrace{\cos\theta}_{\substack{\sin\theta\,\cos\varphi \,:\, \vec{x} \\ \sin\theta\,\sin\varphi \,:\, \vec{y}}} |\Phi_2^\circ|^2\, du\, d^2S$$

Ex: show that the total electric charge is given by

$$Q = -\frac{1}{2\pi} \oint_{\substack{u=u_0 \\ \text{At } \mathscr{I}^+}} \mathrm{Re}\, \Phi_1^0 \, d^2S \qquad \text{for any } u_0.$$

This reconfirms the interpretation of Φ_2^0 & Φ_1^0 at $\mathscr{I}^+$ as capturing the 'radiation' and 'coulombic' information of any given solution. If $\Phi_2^0 = 0$: No energy, momentum or angular momentum carried away $\to$ No EM waves.

At $\mathscr{I}^-$: $u \to v$; so $n^a \to \ell^a$ so, peeling properties reversed;

$$\Phi_0 = F_{ab}\, n^a \ell^b = \frac{\Phi_0^0}{r} + O(\tfrac{1}{r^2})$$

$$\Phi_1^0 = \tfrac{1}{2} F_{ab}\, (n^a \ell^b + m^a \bar{m}^b) = \frac{\Phi_1^0}{r^2} + O(\tfrac{1}{r^3}) \quad \left(\substack{\text{same as} \\ \text{at } \mathscr{I}^+}\right)$$

$$\Phi_2^0 = F_{ab}\, n^a \bar{m}^b = \frac{\Phi_2^0}{r^3} + O(\tfrac{1}{r^4}).$$

so radiative information at $\mathscr{I}^-$ is encoded in $\Phi_0^0(u,v,\varphi)$. No incoming radiation $\Leftrightarrow$ Retarded solution $\Leftrightarrow$ $\Phi_0^0 = 0$ @ $\mathscr{I}^-$.

Asymptotic flatness @ null infinity in GR

Isolation of 'gravitational radiation' in a solution of Einstein's equation: several conceptual and mathematical subtleties.

(i) Again need to go far away from sources. But no natural r-coordinate. Distances defined by g_{ab}, itself the dynamical field!

(ii) What may look like "time-independent" in one coordinate system may appear "wave-like, ondulating" in another because the timelike killing vector you found in a patch is "boost-like" & not a "translation". Example: Levi-civita c-metric. (AA & T. Dray, CMP 79, 581-599 (1981))

→ Led to a lot of confusion about reality of gravitational waves. Einstein had derived the quadrupole formula in the linearized approximation, showing how sources create GWs (1916-1918). But then till 1960's, there was considerable confusion on whether GWs exist in full, non-linear GR.

→ clarified fully by Bondi, Sachs, Newman Penrose and others. This is what I will discuss in the next two lectures. This is the foundation for all current work on GWs.

I.A Einstein-Rosen GWs: Fascinating History

• Einstein 1916: Quadrupole formula showing that general relativity (GR) admits gravitational waves (GWs) in the weak field approximation around Minkowski space. Parallel with Maxwell's theory in striking contrast with Newtonian gravity.

• But then based on his work with Nathan Rosen, in 1936, he sent a paper to Phys. Rev. entitled Do GWs exist? The same day, he wrote to Max Born: *"Together with a young collaborator I arrived at the interesting result that gravitational waves do not exist though they had been assumed to be a certainty in the first approximation. This shows that non-linear gravitational wave field equations tell us more or, rather, limit us more than we had believed up to now."*

Einstein

Rosen

Robertson

Einstein submitted three papers to Phys. Rev. in 1936. Only this paper was sent to a referee. Received a 8 page report (from H.P. Robertson) showing that there was an error, not in the solution itself, but in their conclusion. Einstein and Rosen had curious reactions.

I.A Einstein-Rosen GWs: Final publication

- The paper finally appeared in the proceedings of the Franklin Institute, but in the proofs Einstein reversed the conclusion and changed the title! Nathan continued to believe the original conclusion!!

Journal of the Franklin Institute

Volume 223, Issue 1, January 1937, Pages 43-54

On gravitational waves

A. Einstein, N. Rosen

https://doi.org/10.1016/S0016-0032(37)90583-0

Abstract

The rigorous solution for cylindrical gravitational waves is given. For the convenience of the reader the theory of gravitational waves and their production, already known in principle, is given in the first part of this paper. After encountering relationships which cast doubt on the existence of rigorous solutions for undulatory gravitational fields, we investigate rigorously the case of cylindrical gravitational waves. It turns out that rigorous solutions exist and that the problem reduces to the usual cylindrical waves in euclidean space.

Lecture 2: Asymptotic flatness @ Null Infinity (Penrose, Geroch, AA)

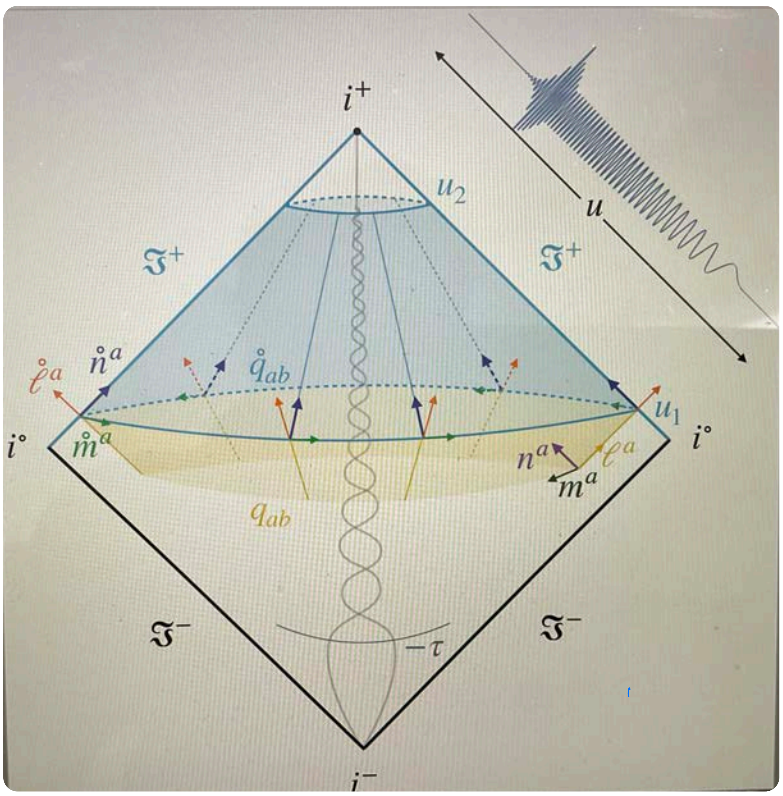
Defn: A space-time $(\tilde{M}, \tilde{g}_{ab})$ is said to be AF@NI if we can attach a boundary $\mathscr{I} = S^2 \times \mathbb{R}$ to M, st on $M = \tilde{M} \cup \mathscr{I}$:

(i) Metric $g_{ab} = \Omega^2 \tilde{g}_{ab}$ (of $-+++$ signature, as $\tilde{g}_{ab}$ is on $\tilde{M}$) and Ω smooth on M; $\Omega = 0$ @ $\mathscr{I}$, and $\nabla_a \Omega \neq 0$ @ $\mathscr{I}$.

(ii) $\underbrace{\Omega^{-2}\, \tilde{T}_{ab}}_{\text{satisfied by physically interesting Maxwell \& scalar fields ; Mink space}}$ $(\equiv (\Omega^{-2}/8\pi G_N)\, \tilde{G}_{ab})$ is smooth @ $\mathscr{I}$.

Definition: Astonishingly simple. No reference to Bondi's expansions, Penrose's null geodesics; coordinate free, yet captures all we need to discuss GWs in full, non-linear GR. (Intuitively $\Omega \sim 1/r$ and $\mathscr{I}^+$ coordinatized by (u, θ, φ)).

In these lectures, I will assume $\tilde{T}_{ab}$ vanishes in a nbd of $\mathscr{I}$ for simplicity. But everything we discuss will go through if $\Omega^{-2}\tilde{T}_{ab}$ has a limit. can have EM radiation (NS-NS coalescence)



(compact binary (NS-NS) evolution depicted in the figure.)

Example: Schwarzschild soln (a star for concreteness, but a BH geometry is identical near $\mathscr{I}$). Again, consider $\mathscr{I}^+$ for definiteness

$$u = t - r_*, \quad r_* = r + 2M \ln\left(\frac{r}{2M} - 1\right)$$

$$d\tilde{S}^2_{sch} = -\left(1 - \frac{2M}{r}\right) du^2 - 2\,du\,dr + r^2 d\omega^2$$

$$\Omega := 1/r$$

$$ds^2_{sch} = \Omega^2 d\tilde{S}^2_{sch}$$
$$= -\Omega^2\left(1 - \frac{2M}{r}\right) du^2 + 2\,du\,d\Omega + d\omega^2$$
on $\mathscr{I}^+$ $\hat{=}$ $2\,du\,d\Omega + d\omega^2$ (same as in Mink space)

Note: go to $\mathscr{I}$ along $u = u_0$, $r \to \infty$

(For $\mathscr{I}^-$, use $v = t + r_*$)

Ex: Do all the intermediate calculations. Also, start with the Kerr-Schild form of the metric and show that $\mathscr{I}$ of Mink space is the same as $\mathscr{I}$ of Schwarzschild. Can do this also for Kerr.

Consequences of field equations $\tilde R_{ab}=0$ near $\mathscr{I}$.

- $g_{ab} = \Omega^2 \tilde g_{ab} \Rightarrow \underset{\overset{\|}{0}}{\tilde R_{ab}} = R_{ab} + 2\Omega^{-1}\nabla_a\nabla_b\Omega + (\Omega^{-1}\nabla^m\nabla_m\Omega - 3\Omega^{-2}\nabla^m\Omega\,\nabla_m\Omega)\, g_{ab}$ (1)

  Indices raised using g_{ab}

  Ex: check this. If you need help, see Wald's book, Appendix D

- $\Omega \hat= 0$ (on $\mathscr{I}^+$) and g_{ab} smooth at $\mathscr{I}$ $\Rightarrow$ strong consequences. Set $n^a = g^{ab}\nabla_b\Omega$

  Then; since $g_{ab} = \Omega^2 \tilde g_{ab}$, $\tilde g^{ab} = \Omega^2 g^{ab}$ whence,

  $0 = \tilde R = \Omega^2 R + 6\Omega\, \nabla^m n_m - 12\, n^m n_m$ (2) $\Rightarrow$ $n^m n_m \hat= 0$.

  $\Omega \hat= 0$ and $g^{ab}\nabla_a\Omega = n^a$ $\Rightarrow$ n^a is normal to $\mathscr{I}$ $\Rightarrow$ $\mathscr{I}$ is a null 3-surface

  (If $\Lambda > 0$, $\tilde R > 0 \Rightarrow n^m n_m < 0 \Rightarrow \mathscr{I}^+$ is spacelike; If $\Lambda < 0$: $\mathscr{I}^+$ is TL.)

- Now $\Omega \times (1) = 0$ $\Rightarrow$ $\nabla_a n_b \propto g_{ab}$ @ $\mathscr{I}$ $\Longleftrightarrow$ $\nabla_a n_b \hat= \frac{1}{4}(\nabla_c n^c)\, g_{ab}$.

  conformal freedom: $\Omega \to \Omega' = \omega\Omega$ where ω smooth, $\neq 0$ on $\mathscr{I}$

  $\Rightarrow$ can always choose ω st $\nabla'_c n'^c \hat= 0$ $\Rightarrow$ $\nabla'_a n'_b \hat= 0$.

  Ex: check this!

- Thus, field eqns near $\mathscr{I}$ (i.e. $\lim \Omega^{-2} \tilde T_{ab}$ exists) imply

  (i) $\mathscr{I}$ is a null 3-surface & (ii) can always choose a conformal factor Ω s.t. $\nabla_a n_b \hat= 0$; (Divergence-free conf. frame).

- Restricted conformal freedom: $\Omega \to \Omega' = \omega\Omega$ s.t. ω is smooth & non-zero @ $\mathscr{I}$, and ($\nabla_a n^a \hat= 0$ & $\nabla'_a n'^a \hat= 0$) $\Longleftrightarrow \mathcal{L}_n \omega \hat= 0$

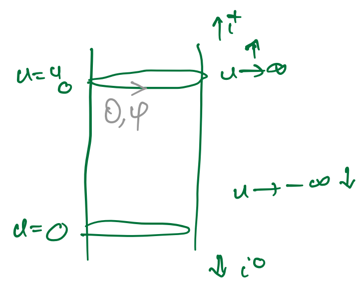
  $\Rightarrow \omega \hat= \omega(\cancel{u}, \theta, \varphi)$; $\Longleftrightarrow \mathcal{L}_n \omega \hat= 0$ since $n^a\partial_a = \frac{\partial}{\partial u}$ $\quad\uparrow$ Exercise.

- In the divergence-free conformal frame:

  $\mathscr{I}^+$ is a cylinder; $\quad n^a \nabla_a u = 1$

  (u, θ, φ) $\quad n^a \nabla_a \theta = 0 \quad n^a \nabla_a \varphi = 0$

  choose a fiducial cross-section



- In the literature, one often further restricts the conformal freedom by demanding that the 'metric' $q_{ab} \hat= \underline{g_{ab}}$ on $\mathscr{I}^+$ $(0,++)$ is a unit 2-sphere metric, i.e. has scalar curvature $=2$

  $q'_{ab} = \omega^2 q_{ab} \Rightarrow R' = \omega^{-2}(R - 2D^2 \ln\omega) = 2$

  $ds'^2 = d\theta^2 + \sin^2\theta\, d\phi^2$

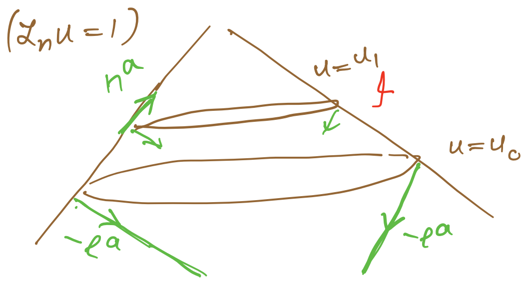
  Sol^{ns} exist: 3-parameter freedom $\omega = \dfrac{1}{\alpha_0 + \alpha_i \underline{r}^i}$; $(-\alpha_0^2 + \alpha_i\alpha^i = -1)$

  $\underline{r}^i = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ unit radial vector

- Advantages & disadvantages of Bondi conf. frames.

Relation to older literature : No conformal completion
Bondi-Sachs & Newman-Unti Asymptotic Expansions

(Asymptotic fall-off imposed on the physical metric in suitable coordinates)



Given such a conformal completion, we can introduce co-ordinates $u, \theta, \varphi, \Omega$ in a nbd of $\mathscr{I}$: Bondi-type expansion

Fix a 2-sphere cross-section C_0 coordinatize it with θ & φ. Define u by $\mathcal{L}_n u \,\hat{=}\, 1$ and $u = u_0$ on C_0. So u is an affine parameter of n^a. We have a family of cross-sections $u = \text{const}$ of $\mathscr{I}$. Let ℓ^a be the other null to these cross-sections, normalized st $n^a \ell^b g_{ab} = -1$. Consider ingoing geodesics generated by $-\ell^a$; $(\ell^a \nabla_a \ell^b = 0)$. Then, introduce u, θ, φ in a nbd of $\mathscr{I}$ using $\mathcal{L}_\ell u = 0$. $\mathcal{L}_\ell \theta = 0$. $\mathcal{L}_\ell \varphi = 0$. set $\Omega = 1/r$

Then, the physical metric $\tilde{g}_{ab}$ has the form:

$$\tilde{g}_{ab}\, d\tilde{x}^a d\tilde{x}^b = -\hat{f}_1\, du^2 - \hat{f}_2\, du\,dr + r^2 (\mathring{q}_{AB} + h_{AB})(dx^A - f^A du)(dx^B - f^B du)$$

Ex.: show this!

Here: $A, B = \theta, \varphi$ and $\mathring{q}_{AB} \equiv$ Unit 2-sphere metric.

$$f_1 = f^{(0)} + \frac{f^{(1)}}{r} + O\left(\frac{1}{r^2}\right); \quad f_2 = -2 + \frac{f_2^{(1)}}{r} + O\left(\frac{1}{r^2}\right); \quad h_{AB} = \frac{\overset{(1)}{h}_{AB}}{r} + O\left(\frac{1}{r^2}\right); \quad f^A = \frac{f^{A(1)}}{r} + O\left(\frac{1}{r^2}\right)$$

Remark : $f_1, f_2, \overset{(1)}{h}_{AB}, f^A$: 7 functions.
Using co-ordinate freedom, can eliminate 1 by further restriction

Newman-Unti choice: $r \to \bar{r}$: Affine parameter of ℓ, so $1 = \tilde{\ell}^a \partial_a \bar{r} = \tilde{g}^{ab} \partial_b u\, \partial_a \bar{r}$: Fixes $f_2 = 1$.

Bondi-Sachs choice: $r \to r'$: "Luminosity distance"

r' : Luminosity distance

determinant of the 2-sphere metric $D = r'^2 \sin^2\theta$

$r, \bar{r}, r'$ have the same asymptotic behavior.

Thus the co-ordinate expansions one finds in the literature can be arrived at starting from conformal completion in a systematic, geometric fashion. The extra input corresponds to fixing the restricted conformal freedom $\Omega \to \omega \Omega$ in a neighborhood of $\mathscr{I}$

Radiation field and Peeling properties in exact GR.

- Fix AF space-time $(\tilde{M}, \tilde{g}_{ab})$ and conf. completion (M, g_{ab}) where $\mathscr{I}^+$ is divergence-free, i.e. $\nabla_a n^a \equiv \nabla_a \nabla^a \Omega \hat{=} 0$

- $$R_{abcd} = C_{abcd} + g_{a[c} S_{d]b} - g_{b[c} S_{d]a} \; ; \quad (S_{ab} := R_{ab} - \tfrac{1}{6} R g_{ab})$$

Field eqns $\Rightarrow 0 = \tilde{S}_{ab} = S_{ab} + 2\Omega^{-1} \nabla_a n_b - \Omega^{-2} n^c n_c g_{ab}$

Multiplying by Ω and taking "curl" we obtain
$\Omega \nabla_{[a} S_{b]c} + C_{abcd} n^d = 0$ } Exercise

$\Rightarrow \quad C_{abcd} n^d \hat{=} 0$

Bianchi Idenity + $S^2 \times R$ topology of $\mathscr{I}$ $\Rightarrow$ $C_{abcd} \hat{=} 0$ } Exercise

$\Rightarrow$ $K_{abcd} := \Omega^{-1} C_{abcd}$ has a smooth limit to $\mathscr{I}$.

Asymptotic Weyl curvature: contrast with Maxwell field.

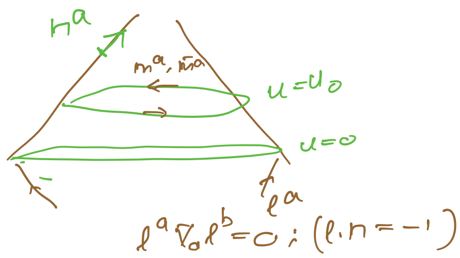
At $\mathscr{I}^+$:

- $K_{bd} :\hat{=} K_{abcd} n^a n^c$, symmetric, Traceless, Transverse: $K_{bd} n^d \hat{=} 0$.

$\Rightarrow$ K_{ac}: 2 components $\equiv \mathbb{E}_{AB}$
: Radiative modes in exact GR
A, B : 1,2 : "Angular components"

comparsion with Maxwell field : $F_b := F_{ab} n^a$, F_b : 2 components $\mathbb{E}_B \longleftrightarrow \Phi_2^0$
($F_b n^b = 0$) 1,2: "Angular components"

Maxwell field: $\mathbb{E}_B$: 1-form on $\mathscr{I}^+$, defined intrinsically. But convenient in practice to write its components in a null tetrad as Φ_2^0.



From $\mathscr{I}$ perspective: we start with a cross-section $u = u_0$, $m^a, \bar{m}^a$ tgt to it : $m^a m^b q_{ab} = 0$, $m^a \bar{m}^b q_{ab} = 1$ and Lie drag $\mathcal{L}_n m^a = 0$ to all of $\mathscr{I}^+$

Then $\Phi_2^0 = F_b \bar{m}^b \equiv \mathbb{E}_B \bar{m}^B$ (depends on choice of $\bar{m}^a$) (spin weight -1)

$\mathrm{Re}\, \Phi_1^0$

For grav. field in exact GR : same procedure used (for convenience)
$\Psi_4^0 = K_{bd} \bar{m}^b \bar{m}^d$: depends on choice of $\bar{m}^b$ ("spin wt" -2)
But convenient in NR, for example. Used heavily in GW literature
Ψ_4^0 called radiation field. : 2 Radiative modes.

Physical null fetrad : Tilde vector fields, $r := \Omega^{-1}$
$\tilde{n}^a = n^a$; $\tilde{\ell}^a = \frac{1}{r^2} \ell^a$; $\tilde{m}^a = \frac{1}{r} m^a$, (as in Maxwell discussion)

Then

$$\Psi_4 := \underset{\text{(physical)}}{\tilde{C}_{abcd}} \tilde{n}^a \tilde{\bar{m}}^b \tilde{n}^c \tilde{\bar{m}}^d = \frac{\Psi_4^0}{r} + O(\tfrac{1}{r^2}) \qquad \text{Radiative part}$$

$$= \frac{K_{ab} \bar{m}^a \bar{m}^b}{r} + O\left(\frac{1}{r^2}\right) \equiv \frac{\mathbb{E}_{AB} \bar{m}^A \bar{m}^B}{r} + O\left(\frac{1}{r^2}\right)$$
(conformally rescaled)

• using field equations, one can find 'potentials' for $K_{ab} \equiv E_{AB}$
These potentials heavily used in NR and waveform models

→ 1st potential: Bondi News $N_{ab} \equiv N_{AB}$ (symmetric, TF, $T \leftrightarrow N_{ab}n^b \hat{=} 0$)

$E_{AB} = \frac{1}{2}\mathcal{L}_n N_{AB} \equiv \mathcal{L}_n(\frac{1}{2} S_{AB})$ $\quad (S_{ab} = R_{ab} - \frac{1}{6} R g_{ab})$

(Note: $N_{ab} = -2(S_{ab} - \rho_{ab})$ $\quad (\mathcal{L}_n \rho_{ab} = 0$ & $\rho_{ab} = q_{ab}$ in a Bondi conf. frame)

$\bar{m}^A \bar{m}^B$: $\quad \Psi_4^o = \dot{N}$ $\qquad N_{AB} = 2(N \bar{m}_A \bar{m}_B + \bar{N} m_A m_B)$

→ 2nd potential: shear $\sigma^o_{ab} \equiv \sigma^o_{AB}$; (symmetric, TF, Transverse)

$N_{AB} = 2\mathcal{L}_n \sigma^o_{AB}$; $\qquad \sigma^o_{AB}$: shear of ℓ^a (conf. space-time)

$N = -\dot{\bar{\sigma}}^o$

$\bar{\sigma}^o \hat{=} -(\nabla_a \ell_b)\bar{m}^a \bar{m}^b \hat{=} -\sigma^o_{AB}\bar{m}^A \bar{m}^B$

$\bar{\sigma}^o = -\sigma^o_{AB}\bar{m}^A \bar{m}^B$

$= \underbrace{\frac{1}{2}(h^o_+ - i h^o_\times)(u,\theta,\varphi)}_{\text{wave form}}$ (Helicity -2)

$= \frac{1}{2} \lim_{r\to\infty} r(\tilde{h}_+ - i\tilde{h}_\times)$

Physical space-time)

$$\boxed{\Psi_4^o = -\ddot{\bar{\sigma}}^o \quad : \text{spin wt 2 because of } \bar{m}^a \bar{m}^b}$$

Ψ_4^o: curvature (Weyl); $\ddot{\bar{\sigma}}^o$: Metric

Coulombic Part

Maxwell:

$\mathrm{Re}\,\Phi_1^o = (F_{ab} n^a)\ell^b$
$= E_b \ell^b$ } charge

physical space:

$\mathrm{Re}\,\tilde{\Phi}_1 = \tilde{F}_{ab}\tilde{n}^a \tilde{\ell}^b$

$\approx \frac{\Phi_1^o}{r^2} + O(\frac{1}{r^3})$

static charge: $\mathrm{Re}\,\tilde{\Phi}_1 = -\frac{1}{2}\frac{q}{r^2}$

Full nonlinear GR:

• $\mathrm{Re}\,\Psi_2^o = (K_{abcd} n^a n^c)\ell^b \ell^d$
$= E_{bd}\ell^b \ell^d$ } BMS energy-momentum

Physical space:

$\mathrm{Re}\,\Psi_2 \equiv \tilde{C}_{abcd}\tilde{n}^a \tilde{n}^c \tilde{\ell}^b \tilde{\ell}^d$

$\approx \frac{\mathrm{Re}\,\Psi_2^o}{r^3} + O(\frac{1}{r^4})$

Kerr : $\mathrm{Re}\,\Psi_2 = -\frac{GM}{r^3}$

• $\Psi_1^o = K_{abcd}\ell^a n^b \ell^c m^d$ } Angular momentum @ $\mathscr{I}^+$

Physical space :

$\Psi_1 \equiv \tilde{C}_{abcd}\tilde{\ell}^a \tilde{n}^b \tilde{\ell}^c \tilde{m}^d$

$\approx \frac{\Psi_1^o}{r^4} + O(\frac{1}{r^5})$

Kerr: $\Psi_1 = \frac{3i}{2r^4}\sin\theta\, G J$

Supplement

Some commonly asked questions.

• Question about "electric and magnetic parts" w.r.t n^a

$\mathbb{E}_b := F_{ab}\, n^a$ $\qquad$ $\mathbb{B}_b := {}^*F_{ab}\, n^a = \frac{1}{2}\epsilon_{ab}{}^{cd}\, n^a F_{cd}$

$\mathbb{E}_b \bar{m}^b \equiv F_{ab}\, n^a \bar{m}^b = \Phi_2^0$ $\qquad$ $\mathbb{B}_b \bar{m}^b = i\, n^c \bar{m}^d F_{cd} = i\,\Phi_2^0$

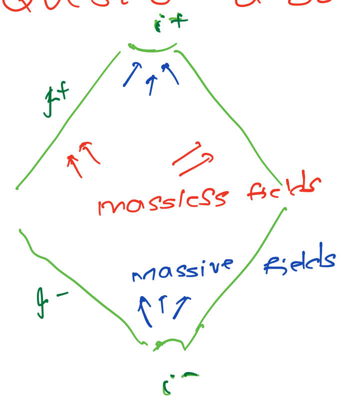
$\mathbb{E}_b \ell^b \equiv F_{ab}\, n^a \ell^b = 2\,\mathrm{Re}\,\Phi_1^0$ $\qquad$ $\mathbb{B}_b \ell^b = 2\cdot \mathrm{Im}\,\Phi_1^0$

Because n^a is null, E_b & $\mathbb{B}_b$ share 2 of 3 components : Φ_2^0

This is in striking contrast with the usual electric and magnetic fields $E_b = F_{ab}\, t^a$ and $B_b = {}^*F_{ab}\, t^a$, with t^a a unit timelike vector, $\perp$ to a spacelike surface. These E_a and B_a are independent.

In the gravitational case, situation is completely parallel; Ψ_4^0 can be extracted as both, $\mathbb{E}_{bd}\, \bar{m}^b \bar{m}^d$ and $\mathbb{B}_{bd}\, \bar{m}^b \bar{m}^d$ $= {}^*K_{abcd}\, n^a n^c \bar{m}^b \bar{m}^d$. $\mathbb{E}_{ab}$ & $\mathbb{B}_{ab} \longleftrightarrow (\Psi_4^0,\ \Psi_3^0,\ \mathrm{Im}\,\Psi_2^0)$ only!

• Question about massive versus massless particles



$\mathscr{I}^\pm$ is the proper arena for discussing radiation, i.e. massless fields

Massive fields (such as neutrinos or scalar fields) are not registered on $\mathscr{I}^\pm$. They come from past timelike infinity i^- and go to future timelike infinity i^+. These are often depicted as points. But one can "blow them up" to 3-d spacelike surfaces (hyperboloids). In Minkowski space, each point of the hyperboloid represents the past and future endpoints of timelike geodesics.

• Question about $K_{abcd} = \Omega^{-1} C_{abcd}$: How can it have a well-defined limit to $\mathscr{I}$ when Ω^{-1} blows up (becomes infinite) there?

$\Omega^{-1} C_{abcd}$ has a well-defined limit because C_{abcd} is smooth and vanishes at $\mathscr{I}$. Consider a smooth function f on the completed manifold M. If $f \hat{=} 0$, the Taylor expansion of f around $\mathscr{I}$ is

$$f = \Omega f_1 + \frac{\Omega^2}{2!} f_2 + \cdots \quad \text{where } f_n = \left.\frac{d^n f}{d\Omega^n}\right|_{\Omega=0}. \quad \text{So } \lim_{\Omega\to 0} \Omega^{-1} f = f_1 + \frac{\Omega}{2} f_2 + \cdots = f_1(\Omega=0, u, \theta, \varphi).$$

Lecture 2 References:

Sections I and II of AA, in the GR Centennial volume, edited by Beiri & Yau; https://arxiv.org/pdf/1409.1800

Section II (parts 8-11) of R. Penrose, Proc. R. Soc (London) 284, 159-203 (1965)

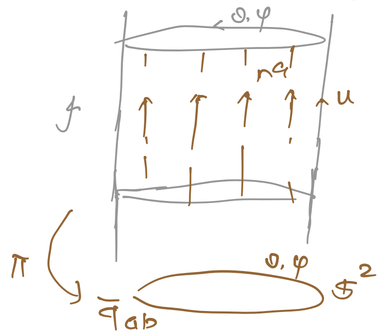
http://igpg.gravity.psu.edu/research/asymquant-book.pdf Pages 44-53

Lecture #3 : Asymptotic symmetries.
The Bondi-Metzner-Sachs (BMS) group

- scalar fields in Mink. space : symmetry group: Poincaré because it preserves the universal kinematical structure shared by all solns to field equation, (eg $\eta^{ab}\nabla_a\nabla_b\phi - \mu^2\phi = 0$) Energy, momentum, angular momentum refer to killing vectors; infinitesimal generators of Poincaré transformations.

- GR: space-time varies from one solution of Einstein's Eq. to another. general solution $\tilde{g}_{ab}$ has no symmetries, ie. killing vectors. But for Asymptotically flat space-times, $\mathscr{I}^{\pm}$ provides a universal background/arena to extract physics.

- Given a conformal completion (M, g_{ab}) of $\tilde{g}_{ab}$, $\mathscr{I}^{\pm} = S^2 \times \mathbb{R}$, $n^a := g^{ab}\nabla_b\Omega$ null normal, $q_{ab} = \underleftarrow{g_{ab}}$ satisfies



$$\mathcal{L}_n q_{ab} = 0 \quad \& \quad q_{ab}n^b \hat{=} 0$$

$q_{ab}: (0++)$ so effectively a metric on S^2.

$\Omega \to \Omega' = \omega\Omega$, then $q'_{ab} = \omega^2 q_{ab}$, $n'^a = \omega^{-1}n^a$.

so $\mathscr{I}^{\pm}$ equipped with pairs (q_{ab}, n^b)

$$(q_{ab}, n^b) \approx (\omega^2 q_{ab}, \omega^{-1}n^b) \; ; \quad \mathcal{L}_n \omega \hat{=} 0$$

- Universal structure shared by all AF solutions to E. Eqns

$\mathscr{I} = S^2 \times \mathbb{R}$, $(q_{ab}, n^b) \approx (\omega^2 q_{ab}, \omega^{-1}n^b)$ st $\mathcal{L}_n \omega \hat{=} 0$

satisfying $\quad q_{ab}n^b = 0 \quad , \quad \mathcal{L}_n q_{ab} = 0$. (contrast with dynamical fields like σ_{ab}, K_{abcd}, ... that vary from one space-time to another)

Degenerate metric 0,++ (a 2-sphere metric)

- (Asymptotic) symmetry group: subgroup $\mathcal{B}$ of diffeo group Diff($\mathscr{I}$) that preserves this structure.

- Lie algebra : $\mathfrak{b}$: VFs ξ^a on $\mathscr{I}$ st

$$\mathcal{L}_\xi q_{ab} \propto q_{ab} \; ; \quad \text{say } \mathcal{L}_\xi q_{ab} \hat{=} 2\phi\, q_{ab} \; , \quad \text{then } \mathcal{L}_\xi n^a = -\phi\, n^b$$
with $\mathcal{L}_n\phi = 0$.

- To explore the structure of $\mathfrak{b}$: consider first the symmetry VF $\xi^a = f n^a$

Ex. check: Then $\mathcal{L}_\xi q_{ab} = \mathcal{L}_{fn} q_{ab} = f\cancelto{0}{\mathcal{L}_n q_{ab}} + 2(q_{m(b} D_{a)} f)\, n^m = 0 \Rightarrow \phi = 0$

$$\Rightarrow \mathcal{L}_\xi n^a = 0 \iff -\mathcal{L}_n(f n^a) = 0 \iff \mathcal{L}_n f = 0.$$

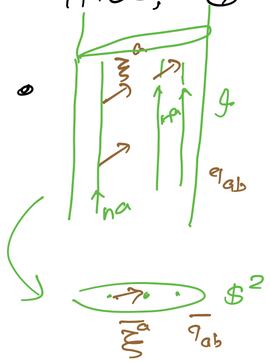
- Thus, if a symmetry VF ξ^a is 'vertical' i.e. $\xi^a = f n^a$
  then $\mathcal{L}_n f = 0$ $\quad (f = f(\cancel{u}, \theta, \varphi))$

These symmetries are called supertranslations, $\xi^a = f n^a \in \mathcal{S}$
constitute an ∞-dim, Abelian Lie-algebra, since

Ex: check: $[f_1 n, f_2 n]^a = \mathcal{L}_{f_1 n} f_2 n^a = \underbrace{(\mathcal{L}_{f_1 n} f_2)}_{=0} n^a - f_2 \underbrace{(\mathcal{L}_n f_1)}_{=0} = 0$.

- Furthermore, given any symmetry VF ξ^a, since $\mathcal{L}_\xi n^a = -\phi n^a$,
  $[\xi, f n]^a = \mathcal{L}_\xi f n^a = (\mathcal{L}_\xi f - \phi) n^a \in \mathcal{S}$ $\quad$ Ex: check $\mathcal{L}_n(\mathcal{L}_\xi f - \phi) = 0$

Thus, $\mathcal{S}$ is a Lie-ideal of the full symmetry Lie algebra $\mathfrak{b}$.



- Any $\xi^a \in \mathfrak{b}$ can be projected unambiguously to a VF $\bar\xi^a$ on the base space $\mathbb{S}^2$

$$\mathcal{L}_\xi q_{ab} = 2\phi\, q_{ab} \Rightarrow \mathcal{L}_{\bar\xi} \bar q_{ab} = 2\bar\phi\, \bar q_{ab}.$$

$\Rightarrow \bar\xi^a$ a conformal KVF of $\bar q_{ab}$ on $\mathbb{S}^2$

Fact: Lie algebra of CKVF of $\mathbb{S}^2$ $\equiv$ Lorentz Lie algebra $\mathcal{L}$.

So, $\underbrace{\mathfrak{b}/\mathcal{S} = \mathbb{L}}_{\text{Lie algebra}}$ $\quad$ or $\quad \mathcal{B} = \mathcal{S} \ltimes \mathcal{L}$

Lie group; semi-direct product

Normal subgroup $\qquad$ quotient

- Recall that Poincaré Lie algebra $\mathfrak{p}$ & Group $\mathcal{P}$ are similar:

$$\mathfrak{p}/\mathbb{T} = \mathbb{L} \qquad\qquad \mathcal{P} = \mathcal{T} \ltimes \mathcal{L}$$

Translations $\qquad$ Lorentz Lie algebra $\qquad$ Normal subgroup

In $\mathfrak{b}$; 4-dim $\mathbb{T}$ is replaced by an infinite dimensional $\mathcal{S}$.

- Big surprise at first: The group is not $\mathcal{P}$ but an infinite dimensional generalization thereof! This comes about because:

AF physical metric: $\underbrace{\tilde g_{ab} = \mathring\eta_{ab}}_{\text{components of } \tilde g_{ab} \text{ in a Cartesian chart of } \mathring\eta_{ab}} + \underbrace{O(\tfrac{1}{r})}_{\text{Remainder}}$

So, $\tilde g_{ab}$ approaches a flat metric $\mathring\eta_{ab}$ in a precise sense.

But if e.g. $t' = t + f(\theta, \varphi)$; $\quad \mathring\eta'_{ab} \longleftrightarrow (t', x, y, z)$

i.e. $\mathring\eta'_{ab}$ is obtained from $\mathring\eta_{ab}$ by an angle dependent translation

Ex: then $\tilde g_{ab} = \mathring\eta'_{ab} + O(\frac{1}{r})$: physical $\tilde g_{ab}$ also approaches $\mathring\eta'_{ab}$.

- Poincaré groups of $\mathring\eta_{ab}$ and $\mathring\eta'_{ab}$ are different. Intuitively, $\mathcal{B}$ is obtained by gluing all these $\mathcal{P}, \mathcal{P}', \ldots$ consistently $\to \infty$ dim.

- $\mathcal{B}$: Asymptotic symmetry group of GR tailored to asymptotic flatness @ null infinity $\longleftrightarrow$ Radiation (GWs, EMWs)

This enlargement came to be appreciated in the particle physics & perturbative treatments of classical & quantum gravity only over the past decade. Conceptually, a key effect of full nonlinear GR.

- Interestingly, the ∞-dim Lie algebra $\mathfrak{S}$ of supertranslations does admit a 4-dim subalgebra $\mathcal{T}$ of translations: unique 4-dim. normal subgroup $\mathcal{T}$ of translations of $\mathcal{B}$. Simplest description: Go to a Bondi Conformal Frame in which q_{ab} is a unit 2-sphere metric. Then: $\xi^a = \alpha(\theta,\varphi)\, n^a \in \mathcal{T}$ if and only if

$$\alpha(\theta,\varphi) = \alpha_0 + \sum_m \alpha_m Y_{1m}(\theta,\varphi)$$

{ Linear combination of first 4 $Y_{\ell m}$ }.

The subspace does not depend on the choice of q_{ab} in BCFs.

- $\Rightarrow$ Notion of energy-momentum (& supermomentum) is well-defined. But angular momentum is more subtle because $\mathcal{B}$ has an ∞-parameter (rather than just 4 as in $\mathcal{P}$) family of Lorentz subgroups $\rightarrow$ supertranslation ambiguity.

Summary

- Asymptotic flatness @ $\mathscr{I}$, needed for GWs in full GR

- The asymptotic symmetry group $\mathcal{B}$: The BMS group preserves the universal structure at $\mathscr{I}$, that is common to all AF space-time. Can also be obtained as:

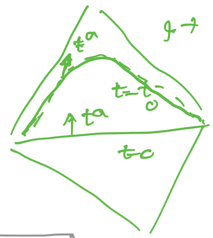
$$\mathcal{B} = \frac{\text{Diff group on } \tilde{M} \text{ that preserves AF boundary conds}}{\text{subgroup of Diffeos that are asymptotically identity}}$$

- $\mathcal{B} = \mathcal{S} \ltimes \mathcal{L}$; $\mathcal{S}$: ∞ dim normal subgroup of supertranslations generated by VFs $\xi^a = f(\theta,\varphi)\, n^a$
under $\Omega \to \omega\Omega$, $n^a \to \omega^{-1} n^a \Rightarrow$ $f \to \omega f$
conformally weighted, wt +1

$\mathcal{S}$ admits a canonical 4-dim subgroup $\mathcal{T}$ of translations generated by $\xi^a = \alpha(\theta,\varphi)\, n^a$; $\alpha = \alpha_0 + \Sigma \alpha_m Y_{1m}(\theta,\varphi)$ in a Bondi conformal frame $\longleftrightarrow$ q_{ab}: unit, round, 2-sphere metric

For further discussion on enlargement of the Poincare group to the BMS, due to supertranslations, see the Appendix of AA, De Lorenzo & Khera, Gen. Rel. Gravit. Grav. 52, 107 (1-27) (2020); https://arxiv.org/pdf/1906.00913.pdf

BMS 4-momentum & supermomentum : Fluxes & "charges"



- Maxwell theory in Minkowski space:

Energy-momentum: source-free solns.

$$F_K = \int_{t=t_0} T_{ab}\, K^a\, dS^b \xrightarrow{t\to\infty} \int_{\mathscr{I}} \alpha(\theta,\varphi)\, |\Phi_2^0|^2\, du\, d^2S$$

Soln F_{ab} to Maxwell eqn — Translational KVF (Ex: show) — $K^a \to \alpha(\theta,\varphi)\, n^a$ — Radiative part of Maxwell field

- For gravity in full GR, we have asymptotic symmetries at $\mathscr{I}$
$\xi^a = f(\theta,\varphi)\, n^a$ supertranslations & $\xi^a = \underbrace{\alpha(\theta,\varphi)}_{1^{st}\ 4\ Y_{\ell m}(\theta,\varphi)}\, n^a$ for translations

But we do not have a gauge invariant notion of stress-energy tensor T_{ab} for the gravitational field!

- Maxwell theory: We can obtain the same expressions of F_K using Hamiltonian methods: F_K is the Hamiltonian on the Maxwell phase space, generator of the infinitesimal canonical transformation $\underbrace{F_{ab} \longrightarrow \mathcal{L}_K F_{ab}}_{\text{Ex: show this is a canonical transformation.}}$ on the space of solutions to Maxwell eqns.
T_{ab} not used!

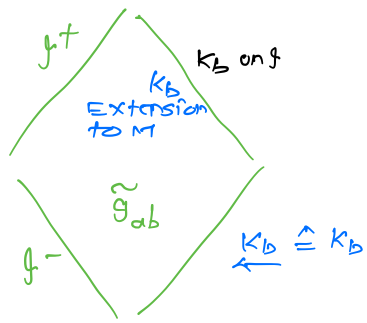
Interestingly we can repeat this procedure for full GR!

- At spatial infinity, this leads to the Arnowitt-Deser-Misner (ADM) expressions of energy-momentum: Total 4-momentum of spacetime including sources & radiation

- We can do the same at null infinity for radiative modes.

Phase space: AA & A. Magon, Comm. Math. Phys. 86, 55-68 (1982).
BMS Hamiltonians: AA & M. Streubel, Proc. R. Soc. (London) A376, 585-607 (1981).

Beautiful mathematical structure associated with geometry at $\mathscr{I}$ leads to the phase space Γ_{Rad} of radiative degrees of freedom, very similar to that in Maxwell & YM theories.



For any asymptotically flat soln $\tilde{g}_{ab}$ of Einstein's equation, at $\mathcal{I}$, conformal completion gives:
(q_{ab}, n^a) at $\mathcal{I}$. "zeroth order" structure to all AF space-times.

$\nabla_a g_{bc} = 0$; ∇_a : derivative operator
connection ($\sim A_a$ in Maxwell)

induces D_a at 3-dim $\mathcal{I}$: $D_a k_b = \nabla_a k_b$ — k_b : extension of k_b on $\mathcal{I}$ to 4-d M.

unambiguous because $\nabla_a n_b \hat{=} 0$

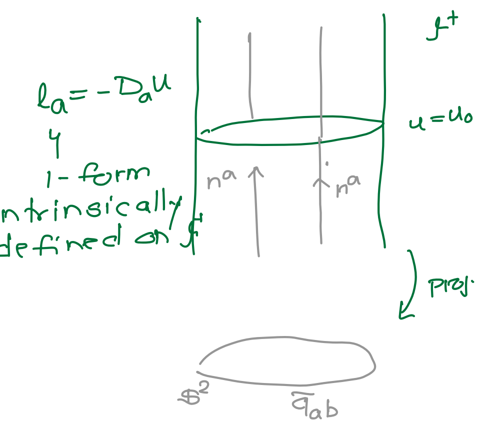
$\nabla_a g_{bc} \hat{=} 0$ & $\nabla_a n^b \hat{=} 0$ $\Rightarrow$ $D_a \underbrace{q_{bc}}_{0,+,+} = 0$ & $D_a n^b = 0$

Because q_{ab} : degenerate, D_a is not unique. It carries "non-universal" information in space-time, i.e. D_a can vary from one physical space-time to another.

Turns out that it carries precisely the radiative information! Non-trivial interplay between physics and geometry; shared by Yang-Mills gauge theories.

$(n^a \nabla_a u = 1)$

New information in D that varies from one solution to another :



Fix any cross-section of $\mathcal{I}$, set $u = u_0$ on it
Then $\mathcal{L}_n u = 1$ $\Rightarrow$ we acquire a 1-parameter-family of cross sections $u = \text{const}$.
normal to the cross-sections : $\ell_a = -D_a u$
$(\ell_a n^a = -1, \quad \ell_a m^a = 0)$

Because $D_a q_{bc} = 0$, the action of D on 'horizontal' h_a $(h_a n^a = 0, \ \mathcal{L}_n h_a = 0)$ is determined, but its action on ℓ_a is not!
$\rightarrow$ can vary from on solution to another.

New information in D $\longleftrightarrow$ shear : $\sigma^{\circ}_{ab} = \text{TF}\ q_a{}^c q_b{}^d \nabla_a \ell_c$.
$= \text{TF}\ \nabla_a \ell_b \equiv \text{TF}\ D_a \ell_b$

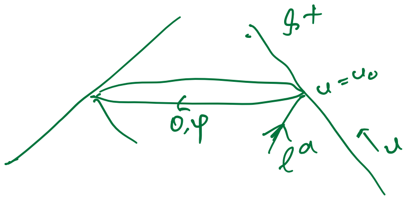
Recall: $\sigma^{\circ} = m^a m^b \sigma^{\circ}_{ab}$: 2-components of Transverse TF σ°_{ab}
$= \underbrace{\tfrac{1}{2}(\dot{h}^{\circ}_{+} + i\dot{h}^{\circ}_{\times})}_{\text{waveform}} \equiv \lim_{\to \mathcal{I}^+} \frac{r}{2} \dot{h}^{tt}_{ab}\ m^a m^b$

Thus, information in D that is not universal/kinematical, i.e. that can vary from one physical spacetime to another is contained precisely in the waveform!

Radiative phase-space : $\Gamma_{Rad} \ni \{ D \}$ on $\mathcal{I}$ $\quad \left(\begin{matrix} D_a n^b = 0 \\ D_a q_{bc} = 0 \end{matrix} \right)$

Subtleties: Possible confusion & Resolution

- **Question:** The null vector field ℓ^a is transverse to $\mathscr{I}$, not tangent. So how can D_a that is an intrinsic derivative operator on $\mathscr{I}$ act on it?



Answer: You are absolutely right. D_a does not know how to act on the vector field ℓ^a. But it knows how to act on co-vector field ℓ_a because ℓ_a is defined intrinsically on $\mathscr{I}$ (i.e. lies in the cotangent space at any point of $\mathscr{I}$). This is at first confusing because this feature (ℓ^a versus ℓ_a) occurs because $\mathscr{I}$ is null.

Recall: The 3-manifold $\mathscr{I}$ is coordinatized by u, θ, φ. So, a triad on $\mathscr{I}$ is $n^a, m^a, \bar{m}^a$ where $n^a\partial_a = \partial_u$ and $m^a\partial_a = \frac{1}{\sqrt{2}}\left(\partial_\theta + \frac{i}{\sin\theta}\partial_\phi\right)$.

The dual covectors are: $\partial_a u$ and $m_a = \frac{1}{\sqrt{2}}(\partial_\theta + i\sin\theta\, \partial_\phi)$
$\ell_a = -\partial_a u$; so $\ell_a n^a = 1$, $\ell_a m^a = \ell_a \bar{m}^a = 0$.

So ℓ_a is in fact a covector defined intrinsically on $\mathscr{I}$.
Note: $n_a = \nabla_a\Omega$ so it's pull-back to $\mathscr{I}$ vanishes $\underleftarrow{n_a} = 0$

Thus, a triad intrinsic to $\mathscr{I}$: $n^a, m^a, \bar{m}^a$
a cotriad intrinsic to $\mathscr{I}$: $\ell_a, m_a, \bar{m}_a$

Therefore the derivative operator D_a knows how to act on ℓ_a : If ℓ_a is any smooth extension to a 4-dimensional neighborhood of $\mathscr{I}$, then $D_a \ell_b = \underleftarrow{\nabla_a \ell_b}$.

- Another clarification: Why is D_a well-defined?
Given any k_a on $\mathscr{I}$, extend it to a neighborhood of $\mathscr{I}$ in M.
If K_a and K'_a are two extensions, i.e. $k_a \hat{=} \underleftarrow{K_a}$ and also $k_a \hat{=} \underleftarrow{K'_a}$ at $\mathscr{I}$, then $(K'_a - K_a) = \Omega V_a + f n_a$ for some VF V_a and function f since $\Omega \hat{=} 0$ and $\underleftarrow{n_a} = 0$. So

$$\nabla_a(K'_b - K_b) = (\nabla_a\Omega)V_b + \Omega\nabla_a V_b + (\nabla_a f)n_b + f\nabla_a n_b$$

pulling back to $\mathscr{I}$:

$$\underleftarrow{\nabla_a(K'_b - K_b)} \hat{=} 0 \qquad \text{since } \Omega \hat{=} 0,\ \underleftarrow{\nabla_a\Omega} = \underleftarrow{n_a} \hat{=} 0,\ \underleftarrow{\nabla_a n_b} \hat{=} 0$$

Hence, $\underleftarrow{\nabla_a K'_b} \hat{=} \underleftarrow{\nabla_a K_b} \hat{=} D_a k_b$: unambiguous.

Thus:

1. At $\mathscr{I}$ we have the symmetry group $\mathcal{B}$ (BMS group) that preserves the universal/kinematical structure common to all AF space-times.

2. Fundamental dynamical field @ $\mathscr{I}^+$: D; captures radiative content of space-time. Through its curvature D determines precisely $\underset{\leftarrow}{K}_{ab} \equiv \mathbb{E}_{ab}$ and hence, N, Ψ_4°, Ψ_3°, $\mathrm{Im}\,\Psi_2^\circ$!

Recall: $N_{ab} = 2\,\mathcal{L}_n \sigma^\circ_{ab} \Rightarrow N = -\dot{\bar{\sigma}}^\circ$ $\qquad N_{ab} = 2(N m_a m_b + \bar{N}\bar{m}_a\bar{m}_b)$

$\underset{\leftarrow}{K}_{ab} \equiv \mathbb{E}_{ab} = \frac{1}{2}\mathcal{L}_n N_{ab} \Rightarrow \Psi_4^\circ = m^a m^b \mathbb{E}_{ab} = \dot{N} = -\ddot{\bar{\sigma}}^\circ$

σ° or $(h_+^\circ + i h_\times^\circ)$ heavily used for waveforms.
Invariant content of σ° is D !

In stationary space-times. D 'trivial'; i.e. completely determined by $q_{ab} \Longleftrightarrow$ we can choose $u=$ const cross-sections of $\mathscr{I}^\pm$ such that

$$\sigma^\circ_{ab} = TF\; q_a{}^c q_b{}^c D_c\, d = 0$$

$$\Rightarrow N_{ab} = 0 \Rightarrow \mathbb{E}_{ab} = 0 \Longleftrightarrow \Psi_4^\circ = 0.$$

In radiative space-times D 'non-trivial'; $N_{ab} \neq 0$, $\Psi_4^\circ \neq 0$
These are the space-times we are interested for GWs.

Using the BMS symmetries $\xi^a \in \mathcal{b}$ and the radiative phase space we can compute fluxes $\mathcal{F}_\xi$ carried by GWs across $\mathscr{I}$: Hamiltonians generating the infinitesimal canonical transformation: $D_a \to D_a + \epsilon\,\mathcal{L}_\xi D_a$

i.e. $D_a k_b \to D_a k_b + \epsilon\,[\mathcal{L}_\xi (D_a k_b) - D_a \mathcal{L}_\xi k_b]$ $\quad \forall k_b$ on $\mathscr{I}$.

(Recall Maxwell theory or Yang Mills or scalar fields in Mink. space.)

$$\mathcal{F}_{\xi^a = f n^a} = \frac{1}{32\pi G}\int_{\mathscr{I}} du\, d^2S \;.\; N^{ab}\left(f(\theta,\varphi)\, N_{ab} + 2 D_a D_b f\right)$$

For translations, $\xi^a = \alpha n^a$, $\quad$ using a Bondi conformal frame for simplicity.

$\alpha = \alpha_0 + \alpha_m Y_{1m}(\theta,\varphi)$;

Exercise: $\quad D_a D_b \alpha = (\alpha_m Y_{1m})\, q_{ab}$. since $N^{ab} q_{ab} = 0$;

$$\mathcal{F}_{t^a = \alpha n^a} = \frac{1}{32\pi G}\int_{\mathscr{I}} du\, d^2S\; \alpha(\theta,\varphi)\, N^{ab} N_{ab}$$

Energy flux: $\alpha = 1 \Rightarrow \mathcal{F}_{t^a = n^a} = \int du\, d^2S\, \underbrace{N^{ab} N_{ab}}_{\geq 0}$

Gravitational waves carry positive energy
"They are real; you can heat water with them!" : Bondi

Commonly used formulas in the GW literature:

Expressions given above do not depend on coordinates or tetrads (or psuedo tensors). At $\mathscr{I}^+$, we can unambiguously calculate fluxes of energy, momentum, ... - covariant expressions

But one often introduces coordinates u, θ, φ to make the expressions "explicit" (eg for numerics). One defines an angular derivative "$\eth$" ("eth") that acts on spin-weighted scalars.

Components in the null tetrad of a TF $f(u, \theta, \varphi)$ is said to carry spin weight s if under $m^a \to e^{i\alpha} m^a$, $f \to e^{is\alpha} f$

Spin weights:
$\sigma := -\sigma_{ab} \bar{m}^a \bar{m}^b : -2$ $\quad N := \frac{1}{2} N_{ab} \bar{m}^a \bar{m}^b : 2$, $\quad \Psi_4^0 := K_{ab} \bar{m}^a \bar{m}^b : -2$

If A has spin wt s, then $\eth A := \frac{1}{\sqrt{2}} (\sin\theta)^s \left(\frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)(\sin\theta)^{-s} A$

Adds +1 to spin weight.

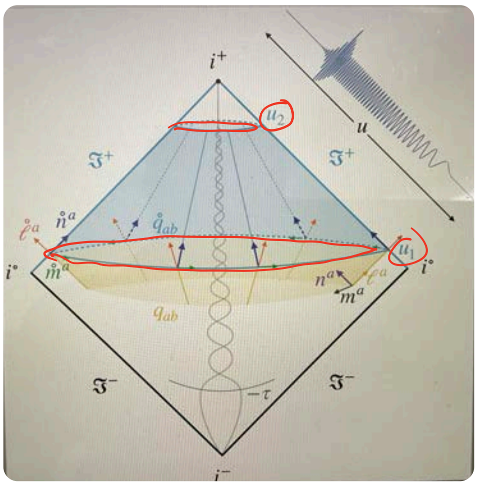
$$\mathcal{F}_{f n^a} = \frac{1}{4\pi G} \int_{\mathscr{I}} du\, d^2S\; f(\theta,\varphi) \left[|\dot{\sigma}^\circ|^2 - \mathrm{Re}\left(\bar{\eth}^2 \dot{\bar{\sigma}}^\circ\right)\right](u,\theta,\varphi)$$

$$\mathcal{F}_{\alpha n} = \frac{1}{4\pi G} \int_{\mathscr{I}} du\, d^2S\; \alpha(\theta,\varphi)\, |\dot{\sigma}^\circ|^2 (u,\theta,\varphi)$$

Wave form
$\bar{\sigma}^\circ = \frac{1}{2}(h_+ - i h_\times)$

Key property of fluxes: Balance laws

The supermomentum flux is an integral over $\mathscr{I}^+$ of an exact 3 form $F^{(\xi)}_{abc} \equiv 3 \nabla_{[a} Q^{(\xi)}_{bc]}$



Hence, given any 2 cross-sections $u = u_1$ & $u = u_2$

$$\int_{\substack{u_1 \\ \Delta\mathscr{I}}}^{u_2} \underbrace{F^{(\xi)}_{abc}}_{3\text{-form}} = \left(\oint_{u=u_1} - \oint_{u=u_2}\right) \underbrace{Q^{(\xi)}_{bc}}_{2\text{-form}}$$

$$Q_{fn^a}[C] \overset{\shortparallel}{\underset{u=u_0}{}} = \oint_{C\,(u=u_0)} Q^{(\xi)}_{bc} \quad : \ \xi\text{-charge at time } u = u_0$$

supermomentum

$$= -\frac{1}{4\pi G} \oint_C d^2S\; f(\theta,\varphi)\; \mathrm{Re}\left[\Psi_2^0 + \bar{\sigma}^\circ \dot{\sigma}^\circ\right]$$

$\mathrm{Re}\,\Psi_2^\circ = K_{ab}\ell^a \ell^b$
$\bar{\sigma}^\circ = \frac{1}{2}(h_+ - i h_\times)$

$\equiv$ supermomentum at retarded time $u = u_0$.

$$Q_{fn^a}[C_1] - Q_{fn^a}[C_2] = \mathcal{F}_{fn^a}[\Delta\mathscr{I}]$$

infinitely many balance laws !!

Portion of $\mathscr{I}$ bounded by C_1 & C_2

Specializing to translations $\xi^a = \alpha(\theta,\varphi)\, n^a$, we obtain the balance law for Bondi-Sachs 4-momentum.

$$\underbrace{P_{(\alpha)}[C]}_{\substack{\text{4-momentum} \\ \text{at retarded time } u=u_0}} = -\frac{1}{4\pi G} \oint_{C \equiv (u=u_0)} d^2S \;\alpha(\theta,\varphi)\; \mathrm{Re}\left[\Psi_2^0 + \bar{\sigma}^0 \dot{\sigma}^0\right](\theta,\varphi)$$

Positive energy theorems proved in early 1980s (Horowitz-Perry, Reula-Tod, Schoen-Yau) show that $P_{(\alpha)}$ is a timelike vector: Energy $\equiv E[C] = -t^\alpha P_\alpha[C] \geq 0$ and vanishes only if space-time is Minkowski. Theorems assume matter satisfies local energy condition.

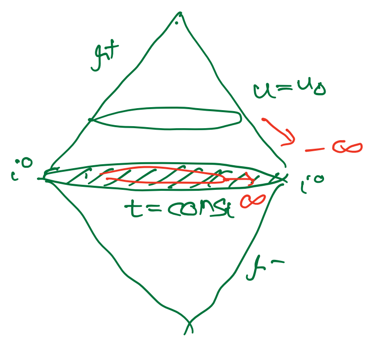
Flux: $$\mathcal{F}_{(\alpha)}[\Delta\mathcal{I}] = \frac{1}{4\pi G} \int_{\Delta\mathcal{I}} du\, d^2S\; \alpha(\theta,\varphi)\, |\dot{\sigma}^0|^2 (u,\theta,\varphi).$$

Balance law. $$P_{(\alpha)}[C_1] - P_{(\alpha)}[C_2] = \mathcal{F}_{(\alpha)}[\Delta\mathcal{I}]$$

special case of supermomentum

Non-trivial agreement:

AA & A. Magnon-Ashtekar, Phys. Rev. Lett. 43, 181-184 (1978)



$$\lim_{u_0 \to -\infty} P_\alpha^{(\text{Bondi})} = P_\alpha^{(\text{ADM})}.$$ (as 4-vectors)

Thus:

- ADM 4-momentum at i^0: Total energy momentum of the physical space-time, including matter & GWs
- Bondi 4-momentum at $u=u_0$: Energy-momentum left over after allowing for radiation to carry away energy-momentum from $u=-\infty$ to $u=u_0$.
- GWs (& EM waves) carry away positive energy but remaining energy at $u=u_0$ is still positive.

One of deepest interplay between geometry & physics!

Summary of the material presented for Lecture 4:
Section. III of AA, in the GR Centennial volume, edited by Beiri & Yau; https://arxiv.org/pdf/1409.1800

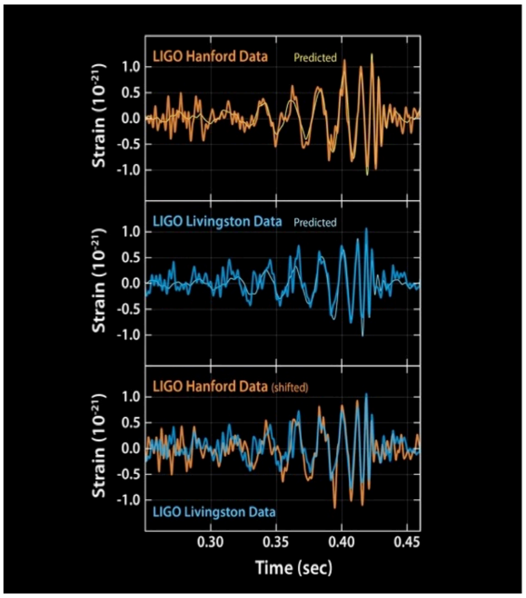
PART II: Balance laws as diagnostic tools for waveforms

In the last two lectures we turn to a concrete application. Recall the plan of the mini- course

(i) Part I: Conceptual and Mathematical issues associated with gravitational waves (GWs) in full, non-linear general relativity. They will thus complement other lectures on approximation methods and numerical relativity by providing the concepts and mathematical notions they use; and,

(ii) Part II: How these results in exact general relativity can be used as diagnostic tools to test the accuracy of model waveforms. Normally one uses numerical simulations to evaluate the accuracy but there are regions of parameter space where numerical simulations are sparse. The diagnostic tests come from identities that must be satisfied in exact GR. Therefore, one can use them to test accuracy of candidate waveforms and suggest directions for improvements in all regions of the parameter space. Furthermore, the balance laws can be used to test accuracy of NR waveforms themselves.

GW 150924



Lecture # 5: Waveforms

Spectacular discoveries of CBCs by GW observatories was possible because of matched filtering method:

① Theoretical waveforms for CBCs (parameterized by initial masses, spins : intrinsic parameters) calculated assuming GR : $\underbrace{\frac{1}{2}(\mathring{h}_+ - i\,\mathring{h}_\times) \equiv \mathring{\sigma}}_{\text{strain}}$

② observed waveform at the detector

Matched-filtering enables one to dig out signal from noise for the discovery and then to determine the intrinsic parameters, the extrinsic parameters (distance & sky location of the binary), and inferred observables : final mass, spin, ... for astrophysics.

⇒ Need theoretical predictions for $\mathring{h}_+ - i\,\mathring{h}_\times \equiv \mathring{\sigma}$ or Ψ_4° @ $\mathscr{I}^+$. to sufficient accuracy.

But in GR we cannot solve the 2-body problem analytically: 2 compact objects — 2 BHs (BBH), or 2 Neutron stars (NS-NS) or a BH and a neutron stars (BH-NS) - spiral in from far and merge.

So we have to use a combination of analytical approximations, and numerical methods. Three stages of evolution :

Inspiral
Long phase in which the two bodies, initially very far, spiral in : Post-Newtonian (PN) . Approximation : Perturbation Theory in (v/c) starting with Newtonian orbits

Merger
Highly non-linear violent phase. Need full GR Numerical Relativity (NR) . A single horizon forms, Huge luminosity

Quasi-normal ringing
single BH surrounded by radiation. Quickly settles down : Perturbation theory around Kerr.

3 Main Avenues to produce the waveforms :

EOB : Effective One Body Approximation : PN + NR input.

Phenom : Phenomenological interpolation : faster IMR PHENOM

surrogate : Extrapolations of NR waveforms filling-in the parameter space.

WAVEFORMS

This is a brief summary of procedures used to create waveforms using PN methods and numerical simulations of Einstein's equations, emphasizing the conceptual aspects and key assumptions and approximations. This is only a bird's eye view addressed to mathematical physicists and therefore glosses over many astute steps and novel techniques that have been used to make nontrivial advances. (This material is based on joint work with De Lorenzo and Khera). The account is not up to date. Nonetheless, this material will enable students to appreciate why non-trivial checks on waveforms are needed and how this purpose is served by the balance laws we discussed in these lectures.

The main focus of the community has been on the part of Compact Binary Coalescence (CBC) that is directly relevant to the sensitivity band of the current gravitational wave detectors. This translates to ~100 quasi-circular orbits where dynamics is expected to be well-modeled by the slow-motion approximation of PN expansions, and the last ~10-15 orbits for which dynamics must incorporate strong field effects of full general relativity. These last orbits are calculated using NR. In principle, one could use NR for the entire process. However, the required computational time and effort would be too large, given that we need to cover an 8 (or greater) dimensional parameter space associated with the binary. That is why a `stitching procedure' is used, where the early waveform comes from the PN analysis and the late waveform from numerical simulations. The result is often referred to as the hybrid wave form. In addition, a number of strategies –the effective one body (EOB) method [1], phenomenological interpolation [2], NR-surrogate models [3] – have been used to enhance the reach of analytical waveforms, and/or to interpolate between parameters used in numerical simulations to create a large bank of waveforms. Thus, while currently there are a few thousand CBC numerical simulations, the data banks contain 100 times as many waveforms. The full bank is used by the LIGO-Virgo collaboration for detection, parameter estimation, and testing GR. (For further details, see e.g., the review articles [4–6] and references therein.)

Various steps in this process involve approximations, guesses based on intuition, and choices that are necessary to resolve ambiguities.

Let us begin with the PN expansion. This is essentially a Taylor series in small velocity –truncated to various v/c orders– which however is not convergent; it is at best an asymptotic series. For example, for luminosity of gravitational waves in the extreme mass limit, the PN expansion starts to deviate significantly from the exact result for $v/c > 0.2$, and the contributions up to $(v/c)^4$ and $(v/c)^5$ terms do so in opposite directions [7]. Consequently, even when one can carry out calculations to a high order, it is not easy to systematically control the truncation errors.

A second issue undergoes the name of Taylor approximants. The post-Newtonian waveforms are obtained starting from the PN expansions of the energy of the system $E(v/c)$ and the flux of radiated energy $F(v/c)$. However, because the procedure involves rational –rather than polynomial– functionals of $E(v/c)$ and $F(v/c)$, there is some freedom in expanding out these quantities to obtain the waveform to a given PN order. Because of this freedom, several different PN waveforms arise at a given order; this is the so-called 'ambiguity in the choice of Taylor approximants.' For unequal masses, this is generally the largest source of errors in the PN waveforms (see, e.g., [7, 8]).

Finally, in the PN literature, there is a fixed background Minkowski space at all orders and the PN solution is assumed to be stationary in the past, before some time $t < -\tau$ [4, 9]. This assumption would seem unreasonably strong to mathematical relativists since for sources for which the initial value problem is well posed in full general relativity, if a solution is stationary in the past in this strong sense, then it is stationary everywhere. However, in the PN strategy the system is non-stationary in the future due to radiation reaction effects and the assumption of past-stationary primarily serves to make various tail terms finite. The viewpoint is that "past-stationarity" is appropriate for real astrophysical sources of gravitational waves which have been formed at a finite instant in the past" [4]. The physical idea behind this strategy is that the two bodies become gravitationally bound at a finite time $t = -\tau$ in the distant past, while being still very far away from one another, and it is argued that the metric perturbation of the background Minkowski space-time can be taken to be stationary before the capture occurs.

In lectures 5 and 6, we will use a much weaker condition, where past stationarity holds in a limiting sense as one goes to past infinity along $\mathscr{I}^+$ and that too only for a certain field. The assumption is mild and expected to hold on physical grounds for CBC (although, not in scattering situations). In particular, it is perfectly compatible with non-stationary solutions in full GR.

In NR we encounter different types of errors. First, there are the truncation errors that are common to all numerical simulations. Second, the wave form is extracted at a large but finite radius, whereas the radiation field becomes truly gauge invariant and unambiguous only at infinity. Therefore, the results inherit error-bars associated with the choice of extraction radius [10]. Third, the waveform is obtained by integrating twice with respect to time the radiation field encoded in the component Ψ_4 of the Weyl tensor. This requires introduction of coordinate systems and null tetrads which become unambiguous only at infinite distance from sources. Finally, although one does have tools to calculate full Ψ_4 (modulo the ambiguities inherent in working at a finite radius), there are numerical errors due to high frequency oscillations which are suppressed if one calculates only the first few (spin-weighted) spherical harmonics because of the 'averaging' involved. Therefore, only the most dominant modes are generally reported in the NR results, rather than the full wave form.

The `stitching procedure' is inherently ambiguous because it involves several choices (see, e.g., [11]). First, one has to decide at what stage in the CMB evolution one stitches the PN and the NR waveforms. Second one must decide which PN order and which T-approximant to use. Third, the PN and the NR waveforms are generally computed using different co-ordinate systems and therefore one has to introduce additional inputs for a meaningful matching. These choices are driven by intuition and guided by past experience rather than clear-cut, unambiguous mathematical physics procedures.

Next, because the PN expansion and NR simulation are based on quite different conceptual frameworks, there are several seemingly ad-hoc elements involved. In PN calculations, the sources are taken to be point particles in Minkowski space. In NR, there is no background Minkowski space and black holes are represented by dynamical horizons (and neutron stars with suitable fluids). In the case of black holes, the individual masses and spins are determined by the horizon geometry. Therefore, for the stitching procedure, one starts with a controlled set of NR initial data (given by the Bowen-York [12] or the Brandt-Brügmann [13] strategy) satisfying constraints of exact GR and evolves. Now, these data contain some 'spurious radiation' which escapes the grid quickly. After this occurs, one re-evaluates the source parameters in the numerical solution and matches them with the source parameters of the PN solution. One then chooses an interval in the time or the frequency domain and evolves both the PN and NR solutions and compares their waveforms. There are several ways to 'measure' the difference between the two waveforms and one minimizes it by tweaking the time of matching, the interval over which the matching is done, and the choice of source parameters in the two schemes.

Conceptually, it is important to note that the matching is done only for the waveform –i.e., for the two asymptotic forms of the metric that capture the radiative modes in the two schemes. In the interior, there is no obvious correspondence between the PN and NR solutions. In particular, there is no simple relation between the 'particle trajectories' representing the black holes, determined by the PN equations, and the dynamical horizons determined by numerical simulations.

These considerations make it clear that even for the ~1% of waveforms in the data banks that are obtained just from PN and NR, there is no systematic way to measure how well they agree with the predictions of exact GR. Inputs that go into the construction of the remaining ~99% of the waveforms are even less driven by fundamental considerations. To mathematical relativists, this can seem shocking. But it is important to note that similar phenomenological considerations and mixture of science and art are heavily used also in other areas of physics, such as QCD.

It is a tribute to the physical intuition and technical ingenuity behind these hybrid waveforms, that the matched-filtering procedure could lead to detections of coalescing binaries.

An Illustrative list of References

[1] A. Buonanno, Y. Pan, H. P. Pfeiffer, M. A. Scheel, L. T.Buchman, and L. E. Kidder, Effective- one-body waveforms calibrated to numerical relativity simulations: Coalescence of nonspin- ning, equal-mass black holes, Phys. Rev. D 79, 124028 (2009).

[2] P. Ajith et al, Inspiral-merger-ringdown waveforms for black-hole binaries with nonprecessing spins, Phys. Rev. Lett. 106, 241101 (2011).

[3] S. E. Field et al, Fast prediction and evaluation of gravitational waveforms using surrogate models, Phys. Rev. X 4, 031006 (2014).

[4] L. Blanchet, Gravitational radiation from post-Newtonian sources and inspiralling compact binaries, Living Rev. Relativity, 17, 2 (2014).

[5] A. Buonanno and B. Sathyaprakash, Sources of gravitational waves, In General Relativity and Gravitation: A Centennial Perspective, edited by A. Ashtekar et al (Cambridge UP, Cambridge, 2015).

[6] L. Blanchet, Analyzing gravitational waves with general relativity, arXiv:1902.09801.

[7] T. Damour, B. Iyer and B. Sathyaprakash, Comparison of search templates for gravitational waves from binary inspiral, Phys. Rev. D 63 044023 (2001).

[8] I. MacDonald et al, Suitability of hybrid gravitational waveforms for unequal-mass binaries, Phys. Rev. D 87, 024009 (2013).

[9] L. Blanchet and T. Damour, Radiative gravitational fields in general relativity I. General structure of the field outside the source, Phil. Trans. R. Soc. (London) A 320, 379-430 (1986).

[10] N. T. Bishop and L. Rezzolla, Extraction of gravitational waves in numerical relativity, L. Living Rev Relativ. 19:2 (2016).

[11] A. Ramos-Buades, S. Husa and G. Pratte Simple procedures to reduce eccentricity of binary black hole simulations, Phys. Rev. D 99 023003 (2019).

[12] J. W. Bowen, J. Rauber and J. W. York, Two black holes with axisymmetric parallel spins: Initial data", Class. Quant. Grav., 1 591-610 (1984).

[13] S. Brandt and B .Brugmann, A simple construction of initial data for multiple black holes, Phys. Rev. Lett., 78 3606-3609 (1997).

Two types of errors:

① Systematic errors: Associated with approximations or "well motivated tricks" used in creating waveform models. Shortcuts are inevitable because we want to cover a 8-dimensional parameter space & computational speed is important. Examples:

(i) Higher ℓ, m modes are often ignored in the (spin-weighted) spherical harmonic decompostion of $(h_+ - i h_\times)$ or Ψ_4^0

(ii) 'Fast' 'practical' way to accommodate precession (i.e. components of the 2 spins $\vec{S}_1$ and $\vec{S}_2$ orthonal to $\vec{L}$.)

(iii) Key inputs from NR are included.

② statistical errors: Associated with fluctuations in the detector noise (a necessary data analysis to extract signal from noise).

So far (for BBH in particular), the systematic errors are smaller than statistical and hence models are adequate for detection, estimation of main parameters and tests of GR (although error-bars could be reduced).

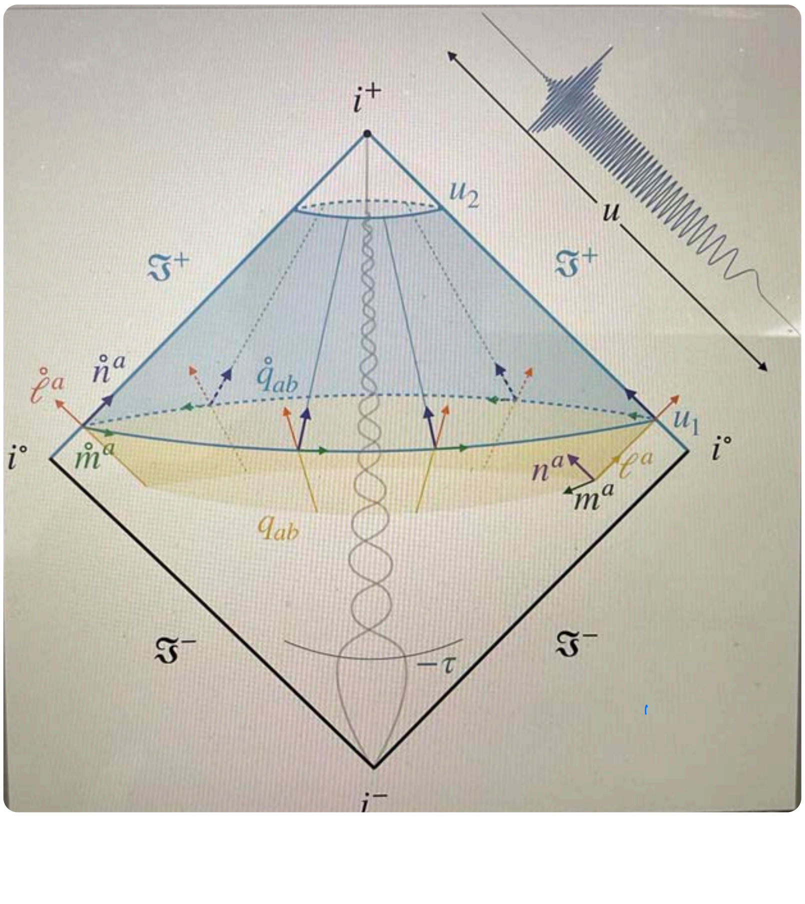
Using the Balance Laws as a Diagnostic Tool

Waveforms from EOB, Phenom and Surrogate models have proved to be invaluable for the impressive detections of the CBCs through gravitational waves. But we are entering an era of abundant event rates (soon, as many as 1000 BBH mergers a year with masses <100 M !) and with LISA and 3g detectors we will achieve a much greater sensitivity over a significantly larger frequency band. For more accurate parameter estimations and tests of GR, one needs quantitative measures of the accuracy of waveforms, relative to exact GR.

- Key problem: We don't know the exact GR waveforms! So in the literature, accuracy tests involve comparisons with NR. But there are big regions in the parameter space where the NR simulations are sparse, so direct comparison is not possible. Also NR results themselves have some errors (e.g., extraction of the wave form at a finite distance; truncation errors; absence of higher harmonics of waveforms).

- Balance laws provide an alternate route that complements NR: Can be used anywhere in the parameter space; and can be used to test NR itself.

- Key feature: Provide an Infinite number of constraints on waveforms, without having to know what the exact GR waveform is ! Whatever the exact waveform is, it must satisfy these supermomentum (and angular momentum) balance laws. Therefore, given any candidate (EOB, Phenom, surrogate, ...) waveform, its violation of these constraints provides and objective measure of how far it is from exact GR, without the need of comparison with the (unknown) exact waveform.



supermomentum balance law:

$\xi^a = f n^a$ supertranslation

$$Q_{fn}[u=-\infty] - Q_{fn}[u=\infty] = F_{fn}[\mathscr{I}]$$

with

$$Q_{fn}[u_0] = -\frac{1}{4\pi G}\oint_{u_0} d^2S \; \underbrace{f(\theta,\varphi)} \; \mathrm{Re}\left[\Psi_2^\circ + \bar{\sigma}^\circ \dot{\sigma}^\circ\right](\theta,\varphi)$$

as $u_0 \to \pm\infty$

$\dot{\sigma}^\circ \sim O\left(\frac{1}{|u|^{1+\epsilon}}\right) \to 0$ as $u \to \pm\infty$ (Finite total energy & ang. momentum radiated across $\mathscr{I}$)

$$F_{fn}[\mathscr{I}] = \frac{1}{4\pi G}\int_{\mathscr{I}^+} dS\, du\, \underbrace{f(\theta,\varphi)}_{\text{Arbitrary}} \left[|\dot{\sigma}^\circ|^2 - \mathrm{Re}\, \eth^2 \dot{\bar{\sigma}}^\circ\right](u,\theta,\varphi)$$

since $f(\theta,\varphi)$ is arbitrary: one constraint for each θ & φ

$\Psi_2^\circ = K_{ab} \ell^a \ell^b$

$$\underbrace{\left[\mathrm{Re}\,\Psi_2^\circ\right]_{u=-\infty}^{u=\infty}(\theta,\varphi)}_{\substack{\text{Determined by initial \& final masses}\\ \text{\& the kick velocity (As shown below)}}} = \frac{1}{4\pi G}\int_{-\infty}^{\infty} du \, \underbrace{\left(|\dot{\sigma}^\circ|^2 - \mathrm{Re}\,\eth^2 \dot{\bar{\sigma}}^\circ\right)(u,\theta,\varphi)}_{\text{provided by the waveform model.}}$$

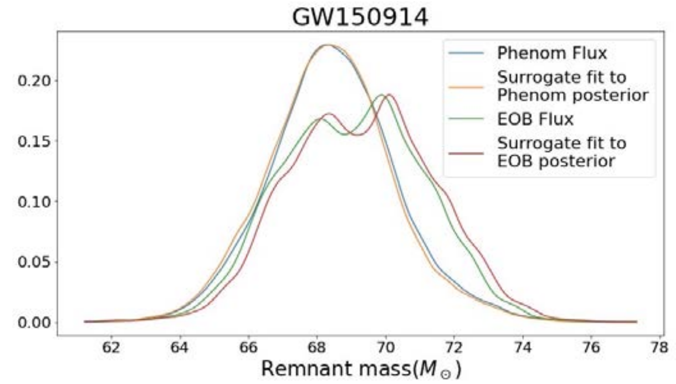
$\bar{\sigma}^\circ = \frac{1}{2}(h_+ - i h_\times)$

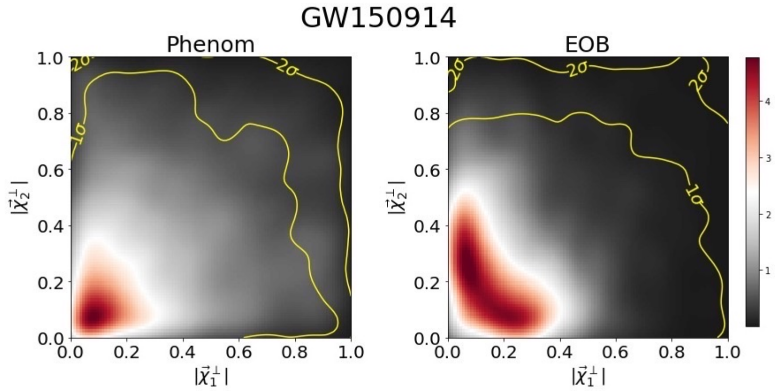
Any candidate waveform must satisfy these infinitely many constraints to desired accuracy.

Let us apply the simplest constraint by integrating the equation with first four $Y_{\ell m}(\theta,\varphi)$: 4-momentum balance law.

$$\underbrace{M_{i+}}_{\text{Inferred observable}} : \text{From } \underbrace{M_{i0}}_{\text{wave form label}} \text{ and } \underbrace{\text{energy-momentum flux}}_{\text{calculated from waveform}}$$

We already learn something about accuracy of waveforms from $Y_{00}(\theta,\varphi)$.





SEOBV3 & IMR Phenom PV2 (Publically available data)

In the posterior probability distribution for the inferred observable M_{fit} (final mass) : Phenom yielded a nice Gaussian. But in EOB: a strange double hump.

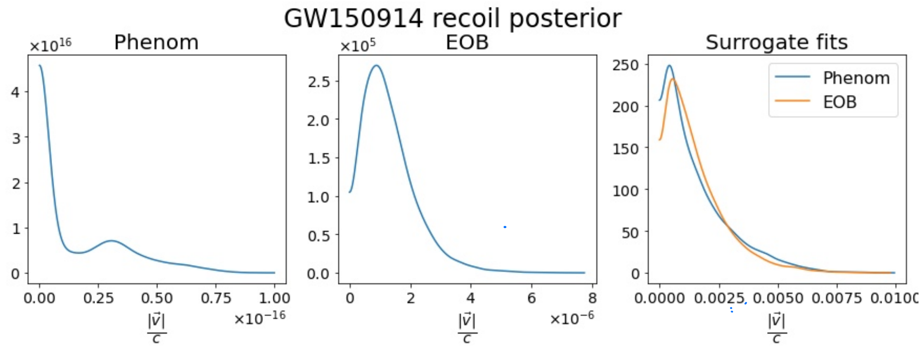
signalled need to re-examine. $\longrightarrow$ Traced to the way EOB treats precession. Leads to a double hump in the posterior probabities of components of two spins $\perp$ to $\vec{L}$. The double hump in the spin-components directly correlated with that in the remnant mass.

2nd example: Posterior probabilities of kick velocity $\vec{v}$:

suppose the 3-momentum carried away is in the x-direction:

$$\vec{P} = P\hat{x}\ ; \quad p \equiv -\frac{1}{4\pi G}\int du\, d^2S\, (\sin\theta\cos\varphi)\, |\dot{\sigma}^{\circ}|^2 \equiv \gamma M_{fit}\, v$$

$$\gamma = 1/\sqrt{1-v^2/c^2}$$



From flux using wave form

NR results

"3-momentum balance law" provids probability distributions.

Balance law is badly violated in both models.

Diagnostics: Higher modes, that are essential to get the correct v were ignored in both models. Have since been incorporated.

Balance law constraint for $\ell \geq 2$.

Usual assumptions made in the CBC analyses:

① $\dot{\sigma}^{\circ}(u,\theta,\varphi) = O\left(\frac{1}{|u|^{1+\epsilon}}\right)$ for some $\epsilon > 0$ as $u \to \pm\infty$

② system is asymptotically stationary as $u \to \pm\infty$
(As $u \to \infty$, Kerr solution; assumption OK
As $u \to -\infty$ usual PN assumption: Quite strong.
one can significantly weaken this requirement ②:

① $\Rightarrow \lim_{\to \mathscr{I}} \mathcal{L}_n K_{ab} \hat{=} 0$ & $\lim_{\to \mathscr{I}} \mathcal{L}_n {}^{*}K_{ab} \hat{=} 0$ i.e. $\dot{\Psi}_4^{\circ}, \dot{\Psi}_3^{\circ}$ & $\dot{\Psi}_2^{\circ} \to 0$ as $u \to \pm\infty$
only require that $\dot{\Psi}_1^{\circ} \to 0$ as $u \to \pm\infty$
Intersting consequence: In the Bondi-frame in which 3-momentum vanishes (at $u = -\infty$ or $u = \infty$), Ψ_2° is real and spherically symmetric.

- However, in general the past and future rest frames are different because gravitational waves carry away 3-momentum:
Black hole kick: Typical NR simulations $\sim \frac{v}{c} \sim$ few 100 km/s.

- suppose the 3-momentum carried away is in the x-direction:
$$\vec{P} = P\,\hat{x}\,; \quad p \equiv -\frac{1}{4\pi G}\int du\, d^2S\,(\sin\theta\cos\varphi)\,|\dot{\sigma}^{\circ}|^2 \equiv \gamma M_{i^+}\, v$$
$$\gamma = 1/\sqrt{1 - v^2/c^2}$$
v: determined by the waveform and the remnant mass M_{i^+}.

- Let us work in the rest frame adapted to i° $(u \to -\infty)$
Then $(\Psi_2^{\circ})(u = -\infty) = G M_{i^{\circ}}$. : spherically symmetric

$\Psi_2^{\circ}(u = +\infty) = G M_{i^+}$ in the Bondi conformal frame in which final BH is @ rest $\vec{P}|_{i^+} = 0$

Two frames related to each other by velocity v. Hence, in the past rest frame $\Psi_2^{\circ}(u = +\infty) = \dfrac{G M_{i^+}}{\gamma^3\left(1 - \frac{v}{c}\sin\theta\cos\varphi\right)^3}$

Thus, $[\Psi_2]_{u=-\infty}^{u=\infty}$ determined by $M_{i^{\circ}}$, M_{i^+} and kick velocity $\vec{v}$

Balance Law:
$$G M_{i^{\circ}} - \frac{G M_{i^+}}{\gamma^3\left(1 - \frac{v}{c}\sin\theta\cos\varphi\right)^3} = \int_{-\infty}^{\infty} du\,\left(|\dot{\sigma}^{\circ}|^2 - \mathrm{Re}\,\eth^2\dot{\bar{\sigma}}^{\circ}\right)(u,\theta,\varphi).$$

Thus all terms in the co no of balance laws

$$\underbrace{\left[\Psi_2^\circ \right]_{u=-\infty}^{u=\infty} (\theta,\varphi)}_{\text{Labels, } M_{i0},\ M_{if}\ \&\ \text{waveform.}} = \int_{-\infty}^{\infty} du \, \underbrace{\left(|\dot{\sigma}^\circ|^2 - \mathrm{Re}\, \eth^2 \dot{\bar{\sigma}}^\circ \right)}_{\text{waveform}} (u,\theta,\varphi).$$

can be computed for each waveform in the catalog! Question: How well does a given labelled waveform satisfy them?

Can decompose the waveform into Y_{lm}'s and obtain a constraint for each (ℓ, m). We already discussed $\ell = 0, 1$. For $\ell = 2$: Dominant modes $m = \pm 2$. In waveform models and NR, one aims at great accuracy for these.

But in the SxS catalog, the $\ell=2, m=0$ mode was poorly represented until recently. Source of error: Believed to be associated with extraction at large (spectral methods) but finite radius. Cauchy characteristic evolution gave accurate results because they extract wave forms at $\mathscr{I}$. But significantly slower. So Khera & SxS collaboration used the finite time version (i.e. $u \in (u_1, u_2)$, rather than $u \in (-\infty, \infty)$) of the balance law to obtain accurate $(2,0)$ mode of the waveform and updated the SxS catalog in 2021.

Mittman, et al, Rev. D 103, 024031 (2021); https://arxiv.org/pdf/2011.01309.pdf (SxS + Penn State)
Mitman et al, Phys. Rev. D 104, 02451 (2021); https://arxiv.org/pdf/2104.07052.pdf (SxS + Penn State)
(Improvement of Recoil velocity in SxS catalog).

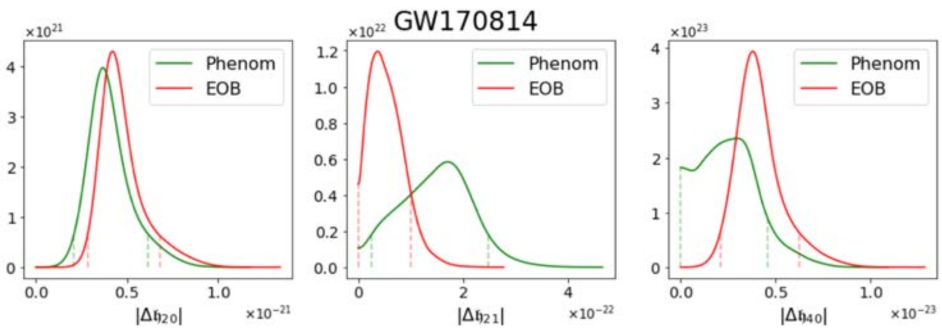
Other illustrative application

Still higher modes in the (spin-weighted) spherical harmonic decomposition.

Khera, Krishnan, AA & Del Lorenzo, Phys. Rev. D 103, 044012 (2021); https://arxiv.org/pdf/2009.06351.pdf

Use gravitational memory as an inferred observable (conceptually just like M_{if} and $\vec{S}_{if}$): obtained by just rewriting the balance law:

$$\left(C_\ell = \tfrac{1}{2}\sqrt{(\ell-1)\ell(\ell+1)(\ell+2)} \right)$$



$$C_\ell (\Delta h)_{\ell m} = -\frac{2G}{D_L^2} \times \left(M_{i0} - \frac{M_{if}}{r^3 (1 - \frac{\vec{v}}{c}\cdot \hat{r})^3} \right)_{\ell m} + \frac{D_L}{2c}\left(\int_{-\infty}^{\infty} dt \, |\dot{\sigma}|^2 \right)_{\ell m}.$$

If two waveform models give statistically different values of this observable, they cannot both be good approximations to GR. : Near parameter values corresponding to this event, one (or both) waveforms need improvement for higher modes.

② Angular momentum balance law: Applications

Khera, AA & Krishnan, https://arxiv.org/pdf/2107.09536.pdf

(Relevant pages included: read the figure captions as well)

For this last topic, the discussion is sketchy because I did not discuss the angular momentum balance law in this course. The purpose of this discussion is only to illustrate ways in which the balance laws can be useful as a diagnostic tool for both waveform models and NR, and can then lead to improvements. In essence each balance law focuses on an aspect of the waveform and serves to bring out limitations that would otherwise be missed.

For the course as a whole, it is interesting to note that the 6 lectures covered a very broad spectrum of ideas that have been developed over 5 decades! The constructions and techniques developed in the 1970s and 1980s still provide foundation for all the forefront theoretical work in GWs. They involve unforeseen and beautiful interplay between geometry and physics. We saw in the last two lectures that, in addition, the older ideas also have a down to earth, practical application as a diagnostic tool to probe the strengths and weaknesses of waveform models vis a via exact GR and to improve them. They can even serve to bring out limitations of NR simulations vis a vis exact GR.improve them. They can even serve to bring out limitations of NR simulations vis a vis exact GR.

This is possible because each balance law enables us to examine the accuracy of the waveform through the lens of a specific observable of exact GR & we have an infinite number of them! This accentuates strengths & limitations of waveforms that are not otherwise apparent.

In the attached paper (Next 4 pages)

See Figures 1, 4 : Limitations of Waveform Models
3, 5 : " " NR simulations

Waveform Models

Fig 1 : Measure of improvements in EOB (from SEOBNRV3 to SEOBNR V4PHM) and Phenom (IMRPhenomPv2 to IMRPhenomXPHM). For Non-precessing systems violation of angular balance law is ~ few percent. Violation is reduced by a factor ~2 for the improved phenom model. Note the "double hump" in the improved EOB model. The model can be improved for the region in the parameter space with $\chi_{eff} < 0.1$ (see Fig 2) This illustrates the effectiveness of balance laws in suggesting directions for improvement of waveforms.

Fig 4 : Need of improvement for precessing systems: particularly the orientation of the final spin. If the balance law held exactly, angle θ would be zero.

Numerical Relativity.

Fig 3 : Top : NRSur7dq4remnant is often used to determine the remnant parameters. Balance law considerations led to an examination of the error estimate provided by the model. The actual errors turned out to be much larger because the estimate did not include errors due to spin-evolution. Although this omission had been noted, the original estimate is widely used.

Fig 5 : The simulation SXS:BBH:1134 is an outlier for which the balance law seemed to be violated by more than 10%! For other simulations violation is < 0.4%. Thus, the balance law raised a red flag. Close examination showed that the reported orbital frequency is erroneous. When corrected, violation is reduced to ~ 0.15%.

Testing waveform models using angular momentum

Neev Khera [1,*] Abhay Ashtekar [1] and Badri Krishnan[2,3,4]

[1] Institute for Gravitation and the Cosmos, Pennsylvania State University, University Park, PA 16802, USA
[2] Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstrasse 38, D-30167 Hannover, Germany
[3] Leibniz Universität Hannover, 30167 Hannover, Germany
[4] Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

The anticipated enhancements in detector sensitivity and the corresponding increase in the number of gravitational wave detections will make it possible to estimate parameters of compact binaries with greater accuracy assuming general relativity(GR), and also to carry out sharper tests of GR itself. Crucial to these procedures are accurate gravitational waveform models. The systematic errors of the models must stay below statistical errors to prevent biases in parameter estimation and to carry out meaningful tests of GR. Comparisons of the models against numerical relativity (NR) waveforms provide an excellent measure of systematic errors. A complementary approach is to use balance laws provided by Einstein's equations to measure faithfulness of a candidate waveform against exact GR. Each balance law focuses on a physical observable and measures the accuracy of the candidate waveform vis a vis that observable. Therefore, this analysis can provide new physical insights into sources of errors. In this paper we focus on the angular momentum balance law, using post-Newtonian theory to calculate the initial angular momentum, surrogate fits to obtain the remnant spin and waveforms from models to calculate the flux. The consistency check provided by the angular momentum balance law brings out the marked improvement in the passage from `IMRPhenomPv2` to `IMRPhenomXPHM` and from `SEOBNRv3` to `SEOBNRv4PHM` and shows that the most recent versions agree quite well with exact GR. For precessing systems, on the other hand, we find that there is room for further improvement, especially for the Phenom models.

I. INTRODUCTION

The next generation of gravitational wave detectors with much higher sensitivity are on the horizon [1–5]. We can expect detection of compact binaries with orders of magnitude higher signal to noise ratio than current measurements. Consequently it will allow unprecedented precision in the tests of general relativity in the highly nonlinear regime. Moreover it will allow high precision parameter estimation of the compact binary. However to carry out these procedures, it is essential to have accurate waveform models whose systematic errors are smaller than the measurement errors.

Gravitational wave observations allow several families of tests of general relativity(GR) [6–8]. Many such tests can be done without waveform models, such as parameterized tests of post-Newtonian (PN) theory [9–13] or tests with the quasinormal ringdown frequencies [14–16]. However these tests rely on the analytic solutions from the perturbative regimes. For testing the highly nonlinear merger regime, waveform models are indispensable. For example one can perform the residual test, where the difference between the data and the best fit waveform obtained from a model is tested for consistency with being purely noise[7, 8]. Some tests can combine many events to have increasing stringency. However it has been shown that accuracy requirements of models also increase for such tests, and that current models may not be sufficiently accurate to perform such tests using detections made so far [17].

Waveform models are created using a diverse range of innovative ideas. However to obtain any model it is necessary to make approximations, and the ensuing systematic errors are unavoidable. A useful way to measure the error is by computing the mismatch of the models against numerical relativity (NR) waveforms using a detectors noise spectrum. If the mismatch $\mathcal{M}$ between NR and the model satisfies $\mathcal{M} \leq 1/\rho^2$, where ρ is the detector signal to noise ratio of an event, then the model will not have significant biases in parameter estimation [18, 19]. Although it has been argued that this sufficient condition can be relaxed in practice [20], nevertheless the mismatch requirement must still scale as $1/\rho^2$. In these analyses one takes NR to be a proxy for the exact GR waveform. Therefore, the accuracy for NR must increase for future detectors as well [21].

On the other hand there are additional tools to measure errors of waveform models from GR: Balance laws. The balance laws don't depend on NR and can thus be used at any point in parameter space, especially where NR simulations are sparse. Moreover the balance laws may provide new insights into sources of errors. Exact GR in asymptotically flat spacetime has a large asymptotic symmetry group: the Bondi-Metzner-Sachs (BMS) group [22, 23]. This group gives rise to infinitely many balance laws [24, 25]. In addition to the more familiar energy, momentum, and the Poincaré angular momentum balance laws, there is an infinite family of supermomentum balance laws. Application of the supermomentum balance law to test waveform systematics was discussed in [26, 27]. The application of the 3-momentum balance

* neevkhera@psu.edu

III. RESULTS

We now apply the methods discussed to waveform models as well as to NR simulations. To test the waveform models across parameter space we select random points in parameter space and check violations of the balance law. We divide our study of the models in two parts: precessing and non-precessing systems. For both these families we restrict the parameter space to a finite compact region. Since we are dealing with binary black holes that are initially in quasicircular orbits, the parameter space is described by the mass ratio q and the dimensionless spins $\vec{\chi}_1, \vec{\chi}_2$. We restrict these parameters to be within range of applicability of `NRSur7dq4`. Additionally, since `NRSur7dq4` only models waveforms for finite time, we would like the `NRSur7dq4` waveforms to be long enough so that we can use PN methods at its start. While `NRSur7dq4` goes up to mass ratio 4, the waveforms start at higher frequencies with increasing mass ratio. Therefore to be able to safely use PN expressions, initially we restrict the mass ratio to $q \leq 2$. This allows us to safely use waveforms starting at 5.8×10^{-3} in dimensionless units. Additionally we also restrict spin magnitudes to be less than 0.8 to be within the training data range of `NRSur7dq4`, as well as the remnant data fit `NRSur7dq4Remnant` that we use.

For the NR simulations we use the publicly available SXS catalog [68] of NR simulations. But we restrict consideration to numerical simulations that lie in the parameter range considered above.

A. Non-Precessing systems

In this section we test satisfaction of the balance law for randomly selected 20,000 non-precessing points in the parameter space. The spins are in the z-direction with χ_1^z and χ_2^z uniformly and independently distributed in the interval $[-0.8, 0.8]$. We obtain the distribution of mass ratio q indirectly from the distribution of masses m_1 and m_2 to replicate commonly chosen priors. We take masses m_1 and m_2 to be independent and uniform, subject to constraints $1/2 < m_1/m_2 < 2$ and $20 < m_1 + m_2 < 160$. Then for each of these points, we will test how well the balance law is satisfied.

We first calculate the spin of the remnant black hole $\vec{\chi}_{\rm bal}$ using the balance law, from Eq. (4). For non-precessing systems, by symmetry we have that $\vec{\chi}_{\rm bal} = a_{\rm bal}\hat{z}$. We can compare this to the remnant spin $\vec{\chi}_{\rm fit} = a_{\rm fit}\hat{z}$ obtained from the fit `NRSur7dq4Remnant`. Mismatch between $\chi_{\rm bal}$ and $\chi_{\rm fit}$ provides us the desired measure of accuracy of the waveform model under consideration. In Fig. 1 we plot the distribution of $a_{\rm bal} - a_{\rm fit}$ across the random points in parameter space. To help identify the errors coming from waveform modelling, we also show an estimate of the errors from the fit. We obtain this by taking the 90% interval of the error estimates provided by `NRSur7dq4Remnant` for the samples of points considered. Similarly we estimate the PN truncation error by using the 90% interval of the distibution of the difference between the 3.5PN and 3PN terms. Although the PN trunction error is not shown in the plot, it is 65% of the fit error, but it does not include the errors from ignoring spin-spin interaction terms.

Measure of improvement

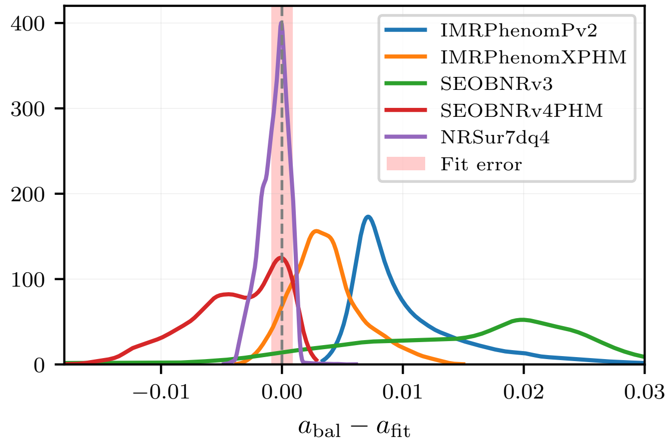


FIG. 1. Non-precessing systems: The distribution of the difference $(a_{\rm bal} - a_{\rm fit})$ between the magnitudes of the remnant spin calculated by using the angular momentum balance law and using the fit `NRSur7dq4Remnant`. The distribution is calculated for different waveform models using the same sample points. The shaded region shows the error estimate of the fit.

Fig. 1 shows that, overall, the agreement between $a_{\rm bal}$ and $a_{\rm fit}$ is of order 10^{-2}. Moreover we see clear evidence for the improvement of `SEOBNRv4PHM` over `SEOBNRv3` and of `IMRPhenomXPHM` over `IMRPhenomPv2`. The surrogate model has the best performance, with all the balance law violation consistent with solely coming from the fit and PN truncation errors. By comparison, although the mismatch is only at a 10^{-2} level for EOB and Phenom, the modelling errors are significantly larger than those coming from the fit and PN truncation errors; thus there is room for further improvement.

Note also that for `SEOBNRv4PHM` the plot has an interesting double hump. We find that these humps are correlated with the effective spin parameter $\chi_{\rm eff}$ defined as

$$\chi_{\rm eff} = \frac{m_1\chi_1^z + m_2\chi_2^z}{m_1 + m_2}\,. \tag{6}$$

The correlation –shown in Fig. 2– brings out the sharp difference between distributions for $\chi_{\rm eff} < -0.1$ and $\chi_{\rm eff} > -0.1$. This illustrates the power of the balance law to identify regions of parameter space where errors are higher, thereby providing guidance for further improvements of the waveform model.

B. Precessing systems

As in Sec. III A, we randomly select 20,000 points in parameter space, but now using precessing systems, and

NR SUR7dq4 Remnant
correcting errorbars

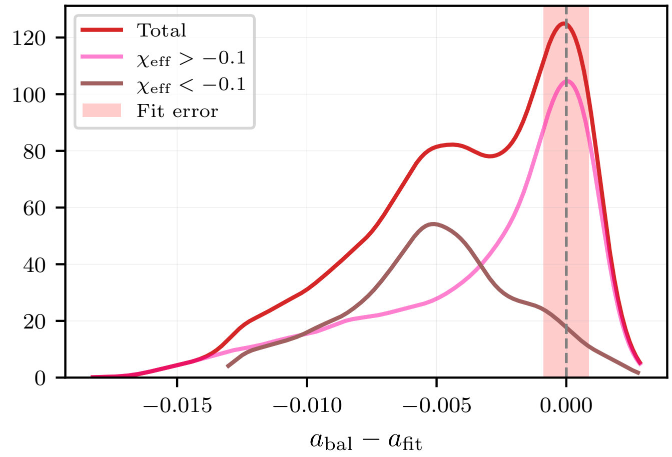


FIG. 2. The distribution of balance law violation for `SEOBNRv4PHM` from Fig. 1. Here we have split the points in parameter space in two, with $\chi_{\rm eff} < -0.1$ and $\chi_{\rm eff} > -0.1$. This split separates the double hump in `SEOBNRv4PHM`, and shows us that the balance law violation is larger for negative $\chi_{\rm eff}$.

evaluate the violation of the angular momentum balance law for them. The spins are sampled independently with an isotropic distribution. The spin magnitude is taken to be uniformly distributed in $[0, 0.8]$. The mass ratio is sampled from the same distribution as in Sec. III A.

The remnant spin is now arbitrarily oriented. Therefore to compare $\vec{\chi}_{\rm bal}$ with $\vec{\chi}_{\rm fit}$, we are led to compare their magnitudes $a_{\rm bal}$ and $a_{\rm fit}$, and also to calculate the angle $\Delta\theta$ between them. However there is a difference in the calculation of error estimates because, as discussed in Sec. II C, for precessing systems the fitting procedure complicated by evolution of spin with time. This is accounted for by using a spin evolution model, which introduces further errors in $a_{\rm fit}$ and $\Delta\theta$. The reported error estimates from the fit `NRSur7dq4Remnant` do not include these errors. Therefore we will estimate these errors by a direct comparison with NR simulations. The NR simulations are taken from the SXS public catalog [68] of NR simulations. We choose quasicircular binary black hole simulations that are long enough to include our choice of starting frequency and have parameters that lie within the range under consideration in this paper. We also drop the first 337 older simulations. We are then left with 672 precessing NR simulations. For these simulations we compute the remnant spin using the fit and compare to the actual NR value. The result is shown in Fig. 3, where we see that the error quoted in `NRSur7dq4Remnant` is much smaller than the actual error. We thus use the 90% interval from these 672 simulations as the error estimate instead. However because the fit is trained against these simulation, the errors might in fact be larger for regions of parameter space with a scarcity of simulations. Nonetheless for the rest of this paper we use these error estimates, keeping in mind that they are not meant to be sharp.

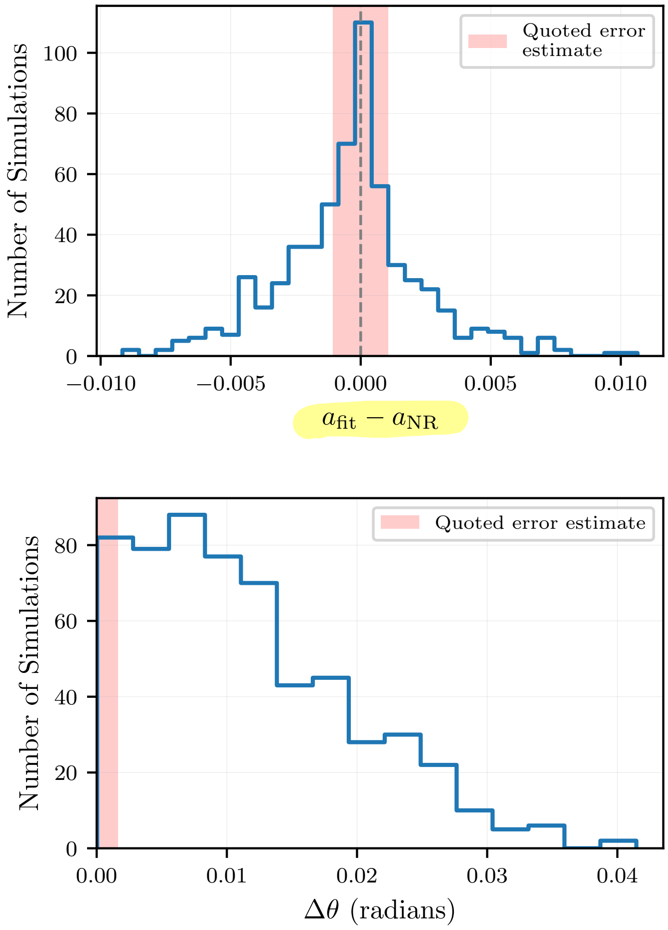


FIG. 3. Comparison of the remnant spin from 672 precessing NR simulations that lie in the parameter range and starting frequency considered in the paper, to the fit `NRSur7dq4Remnant`. The shaded region shows the error estimate provided by the fit model. However as noted in [31], this estimate doesn't include errors from the spin evolution. The upper plot shows the difference in the magnitude of spins, and the lower plot shows the angle between them. We see that for the parameters we consider and for the starting frequency we use, the real errors are much larger than the estimates. We use error estimates obtained from these 672 NR simulations for the rest of the paper.

Using the error estimates discussed above, let us examine the violations of the angular momentum balance law. In Fig. 4 we see the waveform models continue to perform well, albeit with larger errors than in the non-precessing case. For comparisons of the magnitude of the remnant spin, `NRSur7dq4` again has the best performance, and its balance law violations are completely consistent with the error estimates. The PN truncation error is only 9% of the fit error here. The accuracy of the latest EOB and Phenom models, `SEOBNRv4PHM` and `IMRPhenomXPHM`, are very similar to each other. Furthermore, we can clearly see the improvement of these EOB and Phenom models over their older versions. On the other hand, we see different results for the error in the angle in the lower plot of Fig. 4. Here the fit errors are larger. The surrogate

Precessing systems

Correcting orbital frequency in SXS catalog (NR).

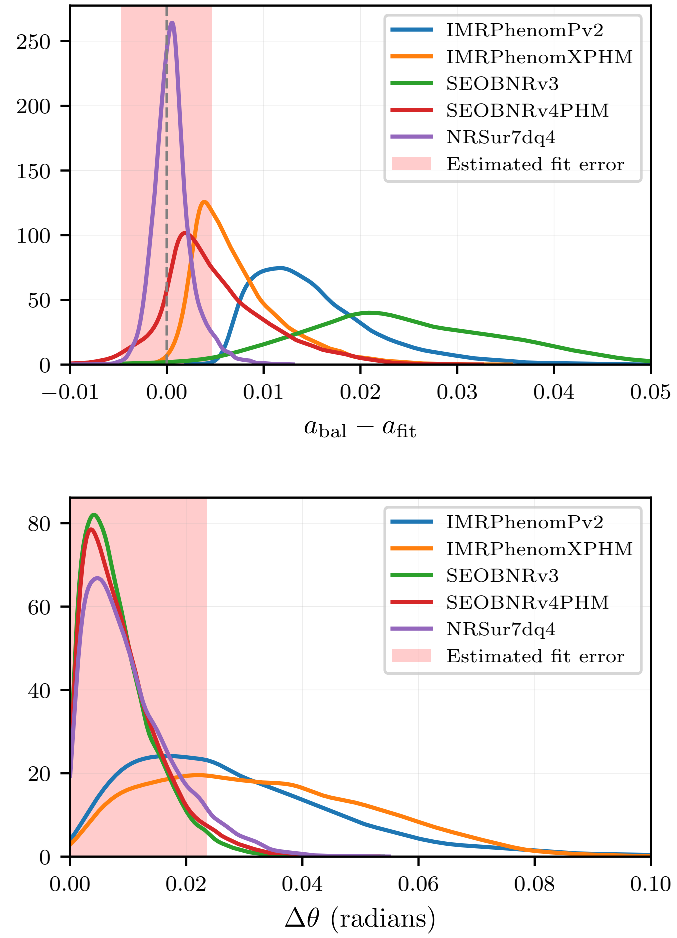


FIG. 4. Precessing systems: The distribution of angular momentum balance law violation across the parameter range considered in the paper, using various waveform models. The upper plot shows the difference between the magnitudes of the remnant spin $a_{\rm bal}$, computed from the balance law, and $a_{\rm fit}$, computed using the fit `NRSur7dq4Remnant`. The lower plot shows the angle $\Delta\theta$ between the remnant spin computed using the two different methods. We also show in the shaded region the error estimate obtained from direct comparison with NR in Fig. 3, as opposed the quoted error estimate in the fit.

and EOB models have violations within the fit errors. The PN truncation error is negligible, only 0.7% of the fit error. However the Phenom models show violations in the angle that are much larger than the errors. Thus, our analysis again provides pointers for further improvement.

C. Lessons from and for NR

We now apply the angular momentum balance law directly to NR simulations and discuss its implications. The procedure is almost identical to the one we used for waveform models, but uses the NR waveform instead of the model waveform. More precisely, each NR simulation

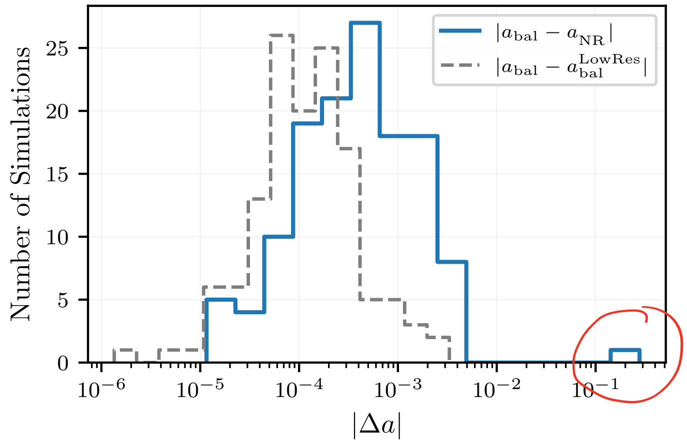


FIG. 5. The violation of angular momentum balance law for the 131 non-precessing numerical simulations described in the text. The solid blue curve shows the difference $a_{\rm bal} - a_{\rm NR}$ between the magnitudes of the remnant spin computed using the balance law, and of the horizon spin. The dashed grey line represents the numerical convergence error, i.e., the difference between the spin magnitudes, $a_{\rm bal}$ and $a_{\rm bal}^{\rm LowRes}$, computed using the highest and a lower resolution NR simulation.

provides us with the waveform to calculate the flux $\vec{\mathcal{F}}$, and is labelled by the masses, spins, orbital frequency and separation of the two progenitors at the starting time. Using these parameters and the 3.5 PN truncation discussed in section II B, we calculate the initial angular momentum $\vec{J}(t_i)$ that is needed in the expression (4) of $\vec{\chi}_{\rm bal}$. For the remnant spin $\vec{\chi}_{\rm NR}$, however, there is a key difference. We do not need the fit since we can directly use the remnant spin computed in the NR simulation at the horizon. The difference $\vec{\chi}_{\rm bal} - \vec{\chi}_{\rm NR}$ measures the violation of the balance law. There is, however, a subtlety: Since the binary system in NR may not be in the same reference frame in numerical simulations as in the frame we use for the PN expression, we must perform a rotation to match the frames. For details see the Appendix.

We use the subset of simulations from the SXS public catalog [68] described in Sec. III B. However we further restrict ourselves to simulations where a lower resolution run is included, allowing us to analyze numerical errors. There are 131 such non-precessing NR simulations and 550 such precessing simulations. For all these simulations we calculate the remnant spin $\vec{\chi}_{\rm bal}$ from Eq. (4) with the highest resolution run available. Then we take the second highest resolution waveform to compute $\vec{\chi}_{\rm bal}^{\rm LowRes}$. Finally, by comparing $\vec{\chi}_{\rm bal}$ to $\vec{\chi}_{\rm bal}^{\rm LowRes}$ we obtain an estimate of the numerical convergence errors, and by comparing $\vec{\chi}_{\rm bal}$ to the horizon spin $\vec{\chi}_{\rm NR}$ we obtain a quantitative measure of the violation of the balance law.

In Fig. 5 the solid (blue) curve shows the violation of the angular momentum balance law for the non-precessing simulations. While the limited number of simulations makes a direct comparison with Fig. 1 difficult, it is clear that overall the errors are manifestly smaller. However there is one outlier simulation `SXS:BBH:1134`

The orbital frequency was erroneous in the meta-data file.
Correcting it brought (Delta A) down from 0.2 to 1.5X10^{-3}!