

Airy-function approach to binary black hole merger waveforms: The fold-caustic diffraction model

José Luis Jaramillo¹ and Badri Krishnan^{2,3}

¹*Institut de Mathématiques de Bourgogne (IMB), UMR 5584, CNRS, Université de Bourgogne Franche-Comté, F-21000 Dijon, France*

²*Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands*

³*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstrasse 38, D-30167 Hannover, Germany*

From numerical simulations of the Einstein equations, and also from gravitational wave observations, the gravitational wave signal from a binary black hole merger is seen to be simple and to possess certain universal features. The simplicity is somewhat surprising given that non-linearities of general relativity are thought to play an important role at the merger. The universal features include an increasing amplitude as we approach the merger, where transition from an oscillatory to a damped regime occurs in a pattern apparently oblivious to the initial conditions. We propose an Airy-function pattern to model the binary black hole (BBH) merger waveform, focusing on accounting for its simplicity and universality. We postulate that the relevant universal features are controlled by a physical mechanism involving: i) a caustic phenomenon in a basic ‘geometric optics’ approximation and, ii) a diffraction over the caustic regularizing its divergence. Universality of caustics and their diffraction patterns account for the observed universal features, as in optical phenomena such as rainbows. This postulate, if true, allows us to borrow mathematical techniques from Singularity (Catastrophe) Theory, in particular Arnol’d-Thom’s theorem, and to understand binary mergers in terms of fold caustics. The diffraction pattern corresponding to the fold-caustic is given in terms of the Airy function, which leads (under a ‘uniform approximation’) to the waveform model written in terms of a parameterized Airy function. The post-merger phase does not share the same features of simplicity and universality, and must be added separately. Nevertheless, our proposal allows a smooth matching of the inspiral and post-merger signals by using the known asymptotics of the Airy function.

I. INTRODUCTION: THE SIMPLICITY OF COMPACT BINARY MERGER WAVEFORMS

A. Compact binary merger waveforms: a status overview

The gravitational wave signal emitted during the merger of two black holes is now calculated routinely by numerical simulations, at least for moderate mass ratios. These numerical simulations solve the Einstein equations with appropriate initial data. The first numerical simulations of binary black hole (BBH) mergers were successfully carried out in 2005 [1–3]. Extensive public catalogs of numerically generated BBH waveform are available [4–6]. On the observational front, signals from binary black hole mergers were first detected by the LIGO and Virgo detectors in 2015 [7]. Since the first detection, the LIGO and Virgo detectors have carried out extensive observational runs with increasing sensitivity, and close to a hundred merger events have now been observed; see e.g. [8, 9] for the most recent event catalogs. The Japanese KAGRA detector has also started to participate in the observing runs [10]. The vast majority of these events are binary black hole mergers, though two binary neutron star mergers and two neutron star - black hole mergers have been detected. The results of the numerical simulations are so far seen to be fully consistent with gravitational wave observations by the LIGO and Virgo detectors (see e.g. [11]). Furthermore numerical relativity results are used in different ways to calibrate analytical waveform models for binary mergers, which are then used to filter gravitational wave data.

These observations and simulations confirm the presence of the three well known regimes of binary black hole mergers, namely the inspiral, merger and ringdown. The early part of

the signal is the inspiral, where the two black holes orbit each other and the orbit gradually shrinks due to energy loss by emission of gravitational radiation. The signal here is quasi-periodic with an increasing amplitude and frequency and can be calculated to a good approximation by post-Newtonian techniques (see e.g. [12, 13]). The late part of the signal is the ringdown, where the final black hole has formed and is approaching its final equilibrium state. Here the signal is a superposition of exponentially damped oscillations, and the frequencies and damping times can be computed within the framework of black hole perturbation theory (see e.g. [14]). Thus far, no consistent approximation scheme can deal with the merger regime and one generally relies on the aforementioned numerical simulations (or analytical models calibrated by these simulations).

A suitable approximation scheme specifically devoted for the merger would be highly desirable, especially in the forthcoming era of third generation gravitational wave detectors. The accuracy requirements on waveforms will become significantly more stringent over the next decade as detectors become more sensitive and can discern finer features of the waveforms. Using less accurate waveforms will likely lead one to draw incorrect conclusions. A clear example is in tests of general relativity and searches for new physics. This problem extends also to binary neutron star mergers. Constraints on the neutron star equation of state rely on measurements of neutron star tidal deformability. These appear as higher order terms in the post-Newtonian approximation, and contribute in the late inspiral regime [15, 16]. Thus, an incorrect merger signal could lead to incorrect measurements of the tidal deformability and eventually to incorrect constraints on the equation of state. Yet another example is the memory

effect. This is a non-linear effect in general relativity which leads to a permanent displacement in the arms of an interferometric detector [17–24]. The contributions to the memory are largest in the merger and thus, more accurate merger models would aid in the quest for direct observations of the memory effect.

B. Simplicity, universality and ‘effective linearity’ of BBH merger waveforms

1. Simplicity and Universality of merger waveforms

Before the first successful numerical simulations of the merger in 2005, it was thought plausible that the waveform could be very complex as we transition from the inspiral to the merger; see e.g. the illustration due to Kip Thorne shown in Fig. 1 of [25]. This figure might suggest that the non-linearities of general relativity could lead to the presence of complicated modulations in the merger signal which might depend sensitively on initial conditions of the binary. However, nature turns out to be much simpler and no such complexity is observed. Indeed, this simplicity is partly responsible for the success of the waveform models mentioned above. On the other hand, the post-merger signal is not universal and depends on the compact object nature of the remnant that is formed by the merger. Thus, for a binary black hole merger, the post-merger signal is determined by the ringing down modes of the remnant black hole. In contrast, for a binary neutron star merger leading initially to the formation of a hyper-massive neutron star, the post-merger signal can be much more complicated. For this reason as we shall see, the post-merger signal needs to be treated very differently.

We should note here that even though the waveform might appear simple, it encodes a non-trivial structure. There is detailed information hidden in the precise manner in which the amplitude and frequency evolve, and modeling this accurately is crucial to extracting reliable information from gravitational wave observations. The fact that the gravitational waveform does not have any wild modulations does not mean that the merger signal does not have important physical information. Rather, it means that the complexities are to be found in small deviations from an underlying simple solution, and most of the difficulty lies in correctly identifying this dominant solution. This simplicity has implicitly helped in the development of several successful complete waveform models which accurately model the inspiral, merger and ringdown phases (see e.g. [26–38]). Using a combination of analytic and numerical methods, these complete waveform models have seen continual improvements over the last decade and especially since the first detections. They are now capable of modeling increasingly non-trivial physical effects such as precession, eccentricity, and can now also incorporate higher angular modes.

In contrast with this approach, we shall not seek to improve waveform models by adding more complexity. Our strategy will rather be to simplify to the largest extent possible. To this end, we identify the following key (‘universal’) qualitative features of the gravitational waveform emitted during a binary

black hole merger waveform:

1. In the inspiral regime, in the context of the post-Newtonian approximation, the signal is a “chirp” signal with increasing amplitude and frequency. Crucially, the post-Newtonian amplitude formally diverges as we approach the merger and the approximation breaks down.
2. Across the merger, the qualitative nature of the signal changes dramatically (‘catastrophically’) from an oscillatory behavior in the pre-merger regime, to a damped behavior post-merger. Thus, the gravitational wave amplitude increases up to the merger, and “shuts-off” very soon after the merger.
3. The post-merger signal is a combination of damped sinusoids for binary black hole mergers. This approximation breaks down as we go backwards in time towards the merger. We note that the signal is possibly much more complicated for mergers involving neutron stars.

The features of the inspiral and merger waveforms are qualitatively similar for mergers involving neutron stars, though the post-merger regime can be dramatically different. These features and the apparent simplicity and universality deserve an explanation. We can also ask: should the merger waveform in modified gravity theories also display this simplicity and universality?

2. Asymptotic Reasoning and Structural Stability

Our strategy is based on two tenets: the Simplicity and Universality of the BBH merger waveforms. These two principles are respectively implemented through the notions of ‘asymptotic reasoning’ and ‘structural stability’.

The first principle is, as mentioned above, to simplify the problem as far as possible to focus on its structural aspects. Such an approach is captured in what can be referred to as “asymptotic reasoning” [39], that aims to uncover the qualitative mechanisms and patterns underlying the observed simplicity [40]. For our purposes, this leads us to disregard effects such as precession, eccentricity etc. which, though physically important, do not modify the essential features of the waveform. These details can be re-introduced at a later stage once the waveform features mentioned above are understood. Moreover, from a more technical perspective, this ‘asymptotic reasoning’ indeed involves an asymptotic limit in terms of an appropriate parameter, in the spirit that asymptotics do capture the relevant underlying patterns, that can (often) be extended beyond the proper asymptotic limit. In our case, such an asymptotic treatment will be provided using an analogy with geometric optics.

The second principle is that of Universality, which refers to the fact that the qualitative features of the merger waveform do not depend on the details on the initial configuration. Universality will be captured and implemented by the notion of structural stability, i.e. stability under small generic perturbations. Here we recognize that mathematical models of natural phenomena are never exact, and thus it is essential for our models

to be robust under small changes of the initial conditions and also in fact, under small changes of the dynamical equations. In particular, under small deformations of general relativity itself, we should expect the above-mentioned essential features of BBH waveforms to persist. Finally, it can be argued that Simplicity and Universality are not independent notions. Indeed, a condition for discussing ‘asymptotic reasoning’ is the realization of some kind of ‘structural stability’ [39]. It will be useful though to methodologically separate these two aspects.

The idea that BBH waveform should be simple and universal is not a new one. In fact, even before the results of the numerical simulations were available, there were several prescient suggestions of such simplicity. A notable example is the early Effective-One-Body model proposed by Damour, Buonanno and collaborators [26, 41]. Their work was based on two fundamental ingredients: i) gravitational radiation leads to a circularization of the orbits of the binary system as it goes through the inspiral regime in a sequence of Keplerian orbits. ii) This “adiabatic” inspiral phase terminates as the system reaches a last-stable orbit. This causes a qualitative transition from the inspiral to a “plunge” phase (followed shortly by the merger and ringdown). This suggests a Universality of the merger signal and is reminiscent of the effacement principle studied by Damour [42], where the internal structure of the two binary components does not appear in the signal at leading order. Apart from this Universality, the Buonanno-Damour proposal suggests that the merger should be simple as well, which in turn contributes to the success of this approach.

A second notable support to Simplicity and Universality is the Close-Limit approximation due to Pullin, Price, Gleiser, Khanna and collaborators [43–45]. It adds an additional ingredient, a sort of ‘effective linearity’, that we expand below.

In summary, simplicity partly explains the success of existing signal models which have been employed in various analyses to date (and are largely responsible for many of the remarkable achievements summarized above). Developing such signal models would have been much more challenging if the late inspiral were to have additional features depending sensitively under small variations of physical parameters.

3. ‘Effective linearity’ of the BBH merger waveform

The Close-Limit approximation mentioned above shows that in the very late inspiral phase, when the two black holes are very close to each other, the spacetime can be surprisingly well approximated as the perturbation of a single black hole. This approach has been used to estimate the recoil velocity of the final black hole [46, 47]. We mention also the Lazarus project [48, 49] which, partly motivated by the success of the Close-Limit approximation, combined the results of numerical relativity with perturbation theory. Again, this approach suggests both universality and simplicity.

The surprising success of the close-limit approximation suggests an “effective linearity” underlying the essential mechanism behind binary black hole merger waveforms. There are other hints as well for linearity in black hole merg-

ers: i) It has been suggested that the linear description of the ringdown phase can be extended all the way back to the merger itself [50–53] (though several open questions remain [54, 55]). ii) Other studies find strong evidence for the presence of linear correlations between geometric fields in the strong field region and fields in the asymptotic wave-zone in a BBH merger [56–63]. These correlations are a dynamical feature difficult to understand without accepting some kind of underlying effective linear ingredient in the emission and propagation of the gravitational degrees of freedom captured in the BBH waveform.

It is important to stress that such effective linearity does not undermine the fundamental non-linear nature of the basic equations and the crucial role of non-linearities in BBH mergers [64]. Indeed, already the Close-Limit approximation makes non-linearities apparent by the need of going to second-order perturbation theory to account for key features of the merger. But, when dealing with the emitted radiation, BBH merger waveforms seem to behave more linearly than expected, as seen above. In this specific setting, the ‘effective linearity’ assumption proposes the effective linear nature of the dominating mechanism underlying the overall qualitative features of the BBH merger waveform. The appropriate background for such linear mechanism to take place should indeed be controlled by the ultimate non-linear equations, its identification requiring a good control of the full BBH dynamics. An avenue to get some insight into this point is discussed in the companion article [65]. Here, we follow a more agnostic approach, adopting the ‘effective linearity’ of the mechanism behind BBH merger waveforms as an assumption.

C. A catastrophe theory approach to BBH mergers: Airy model

Despite the physical insights and far-sighted nature of the Close-Limit and the Effective-one-body models, none of these or other approaches have been able to find a simple functional form for the merger signal. This is also reflected in the fact that the post-Newtonian and Ringdown signals cannot be extended beyond their regimes of validity (cf. however [66]). Thus, the signal amplitude in the post-Newtonian expansion is $(t_c - t)^{-1/4}$ and formally diverges at the coalescence time t_c . Similarly, the ringdown signals diverge exponentially as we go backward in time towards the merger. In essence, the nature of the gravitational wave signal changes from an oscillatory to a damped phase and we do not have yet general model which connects these two asymptotic expansions.

Any analytic description of the merger must necessarily interpolate between these two radically different signal morphologies. In this article we propose such a model for the merger. We shall argue, based on ideas from singularity theory developed by V.I. Arnol’d and R. Thom beginning in the 1960s (building on work by Whitney and others), that the merger waveform should indeed be simple and universal. More concretely, that the Airy function should be the basis for describing the inspiral-merger transition. The arguments presented here do not constitute a proof starting from first

principles, but rather a working hypothesis which needs to be confirmed (or disproved) in further work. In any case, starting with this hypothesis, we shall see that at leading order, all inspiral-merger waveforms can be obtained by a modulated and parameterized Airy function.

From a more systematic perspective, simplicity is addressed in terms of an asymptotic reasoning involving two layers: i) a first one in a ‘geometric optics’ treatment at large frequencies focusing on the formation of caustics, and ii) a second one incorporating wave features, namely diffraction over caustics, at an intermediate level not involving the full (non-linear) wave theory. Universality then relies on the ‘structural stability’ results for stable low-dimensional singularities, following from Arnol’d-Thom theorem in singularity theory. In particular, this theorem classifies the stable low-dimensional caustics, as well as the (universal) diffraction patterns over them. By construction, the proposed framework is stable under small changes in the dynamical equations, and therefore will also hold for small modifications to general relativity.

We will begin our discussion with an outline of the Arnol’d-Thom theory of structurally stable singularities of differentiable mappings, sometimes referred as Catastrophe theory. We shall take two calculations of caustics from optics as prototypical examples for low dimensional mappings. We shall review how the theory of caustics fits into this framework, namely the cylindrical lens and the formation of rainbow. The cylindrical lens is an example of a cusp-catastrophe, while the rainbow is a fold-catastrophe. We shall then argue that a BBH merger can be viewed as a fold-catastrophe, which will lead us directly to the Airy function ansatz. Some simple implications of this ansatz will be considered and finally we shall suggest some directions for future work. Our goal here is to introduce the framework of Singularity theory and Caustics to gravitational wave researchers, and to describe the basic elements of our new model for compact binary mergers. Further developments and details necessary for applying these ideas to gravitational wave data analysis will be discussed elsewhere.

II. CAUSTICS IN GEOMETRIC OPTICS: THE SKELETON FOR A BBH MERGER MODEL

The most evident feature of a BBH merger waveform is that shortly after the merger time t_{merger} , its amplitude rapidly decreases and soon becomes too small to be detectable. Thus, we might say that at times before t_{merger} , the gravitational wave detector is “illuminated” by gravitational waves, and after t_{merger} it is not. Such dramatic changes in illumination are frequently and commonly observed in optics [67] and, in particular, occur when passing through caustics (where the field intensity formally diverges). Moreover, properties such as the diffraction patterns of the electromagnetic wave field near caustics have a universal behavior as we shall see. Let us then take, as a working hypothesis, that the behavior of the BBH waveform near t_{merger} is described by an underlying caustic mechanism. Before proceeding to investigate the validity and implications of this hypothesis, it will be useful to review the standard theory of caustics in geometric optics, and

its description within the framework of Arnol’d-Thom Singularity theory.

A. The rainbow and the lens

In the geometric optics approximation to wave propagation at large frequencies, a propagating wave field corresponds to a family of light rays, and in our case, “gravitational rays”. The determination of the trajectory from a source to a detector of a given ray is obtained, according to Fermat’s principle, from the extremalization of an appropriate function, the phase difference Φ . In the case of monochromatic light this corresponds to the time delay $t = \Phi/\omega$. Given a set of parameters \mathbf{X} (also referred to as “control parameters”) labeling points where a detector could be placed, the extremalization is performed with respect to a set of “state variables” \mathbf{R} that encode the physical features of the source as well as the medium in which the ray propagates. The space of state variables will be denoted S and the space of control variables will be denoted D . The phase difference (‘generating’) function Φ is then a real valued function on $S \times D$.

If a single extremum (‘critical point’) of Φ exists for a given point \mathbf{X}_o in D , then a single ray reaches the detector at \mathbf{X}_o corresponding to a point \mathbf{R}_o in S , that characterizes the ray. In this simple case, the point \mathbf{X}_o in D uniquely determines a point \mathbf{R}_o in S , a single-valued function $\mathbf{R} = \mathbf{R}(\mathbf{X})$ can be defined and no caustics appear. However, if several extremal points of Φ exist for a given \mathbf{X}_o , i.e. that point \mathbf{X}_o is reached by more than one ray, then the mapping $\mathbf{X} \mapsto \mathbf{R}$ is multi-valued. In this situation, we have the possibility that as \mathbf{X}_o is varied, two or more extrema of Φ can coalesce at a given \mathbf{X}_c , that defines a point of the caustic set in the control space D : on one side of the caustic several rays reach \mathbf{X}_o , whereas (some of) these rays disappear at the other side of the caustic. Within the geometric optics approximation, the light intensity diverges in the transition at the caustic. In reality of course (i.e. at finite frequencies away from the geometric optics limit), there is no divergence but there is nevertheless a sharp increase in intensity at the caustic. Such dramatic change from “illumination” to “no-illumination” at a caustic is the mechanism that we intend to explore as providing the ‘skeleton’ for our model of the BBH merger waveform, namely addressing the transition from the inspiral phase to the extinction of the signal.

Before proceeding further, we briefly illustrate the discussion above with two representative examples from optics.

1. Example 1: cylindrical lens.

We first consider an axisymmetric lens, as in Fig. 1 (we follow closely Ref. [68]). Due to the axisymmetry, the problem reduces to two-dimensions and we choose to work in the (x_1, x_2) plane with the incident rays traveling along the x_1 -axis, and the lens is aligned along the x_2 -axis. A bunch of rays parallel to the x_1 axis reaches the lens, each ray arriving at a distinct height a on it. This a represents the “state

variable” \mathbf{R} in this system and the space D of “control parameters” is $\{(x_1, x_2)\}$. Then, rays are focused towards the axis of the lens. We assume that the lens is non-ideal in that the rays do not all meet at the focus, but the rays at a larger a see effectively a longer focal length. It can be shown geometrically that there are three regions A , B and C in the (x_1, x_2) plane: i) only one ray passes through any point in A , ii) three rays pass through each point of B and iii) the curve C with two components meeting in a cusp, the focus of the lens, separates A and B . Thus B is more illuminated than A and the illumination is largest on C , which is the caustic (in particular, the enveloping curve of deflected rays). Within the geometric optics approximation, the illumination in fact diverges at the caustic (where the ray density diverges) and it becomes necessary to invoke the wave nature of light in order to regularize this divergence. In more detail, if we move towards the caustic C from its interior (region B), two of the three rays coincide (“coalesce”) at the caustic, disappearing when passing to the exterior of the caustic (region A). Such drastic ray disappearance corresponds to a singularity at the caustic (in a sense to be specified below), having as a consequence a ‘catastrophic’ fall of the luminosity when passing through the caustic.

Let us revisit the previous discussion in a more systematic manner. The phase difference $\Phi(a; x_1, x_2)$ is a function of the state and control variables. In general this function is multi-valued. For a given ray, the state variable corresponding to a given point in the (x_1, x_2) plane is obtained by extremizing the time delay with respect to the state variable a :

$$\frac{\partial \Phi(a; x_1, x_2)}{\partial a} = 0. \quad (1)$$

In general, multiple rays will pass through a given point so that this equation will have multiple solutions. Solving this equation provides a set of a ’s for a given (x_1, x_2) which in turn corresponds to a hypersurface M in $S \times D = \{(a, x_1, x_2)\}$ space which can “fold” over multiple times. If we define the natural projection from M to D , namely $(a, x_1, x_2) \mapsto (x_1, x_2)$, this projections becomes ‘singular’ precisely at those points where M folds over, their image in D providing precisely the caustic set C (cf. Fig. 2). It is in this sense that the caustic corresponds to a singularity: it is the image of the singular points of the projection from M to D , namely the “folding lines” in the hypersurface M (we illustrate this below).

Alternatively, we can ask ourselves: for which points in D can we write M locally as a function? Specifically (for a fixed x_1), we can ask ourselves for which x_2 ’s in D we can locally write a function $a = a(x_2)$. From the implicit function theorem, this is determined by the differential of the function $\frac{\partial \Phi(a; x_2)}{\partial a}$ defining M in Eq. (1), more specifically this is controlled by the second derivative $\partial^2 \Phi / \partial a^2$ evaluated at points of the solution space M . “Branches” of the manifold M cannot be locally written as a function in those points of D where this second derivative vanishes: such points correspond to the caustic C .

A straightforward calculation shows that the solution hypersurface M is of the form shown in Fig. 2 and the projection onto D is singular at the points when it “folds” over. In more detail, the phase delay function Φ is a quartic polynomial in a

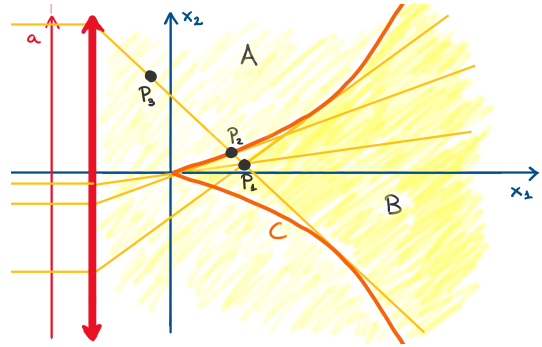


FIG. 1. Cusp caustic for an axisymmetric lens. Each point in the region A has one ray passing through it while each point in C has three rays passing through it. The caustic C separates these two regions.

(cf. e.g. [68])

$$\Phi_{\text{lens}}(a; x_1, x_2) = \frac{(a - x_2)^2 \omega}{2f x_1 c} \left(f - x_1 + \frac{a^2 x_1}{2s^2} \right). \quad (2)$$

where ω and c are the light frequency and speed, respectively, f is focal distance and s is a length scale in the lens. Variables a , x_1 and x_2 are shown in Fig.1 with the coordinates (x_1, x_2) centered at the focal point.

Finding the stationary points of Φ leads to M as a cubic surface folding over itself. Projecting M to the (x_1, x_2) plane, we see that the projection of the singular points is precisely the caustic C , as commented above. Finally, it can be shown that the light intensity I in the geometric optics approximation is proportional to

$$I \sim \left| \frac{\partial^2 \Phi}{\partial a^2} \right|_M^{-1}, \quad (3)$$

which diverges at the caustic. This signals the breakdown of this approximation and it becomes necessary to consider the wave features of light. As an intermediate stage prior to invoking the full Maxwell equations, one can consider the Fresnel diffraction formalism, as we shall shortly discuss.

This example (corresponding to a so-called “cusp caustic” due to the shape of C) has the virtue of being intuitive, but it does not completely capture the transition from “illumination” to “no-illumination”, since rays still reach the outside region. The following example, although less intuitive, is more faithful to the actual mechanism that we propose to be at work in BBH mergers, which we aim to address here.

2. Example 2: The rainbow.

Our second example deals with the physical mechanism of the rainbow (we closely follow Ref. [69]). Let us consider a spherical droplet of radius a and a ray reaching its surface with an incidence angle i . The light is refracted, internally reflected and finally exits the droplet after another refraction. Let the ray enter into the droplet with a refraction angle r . Assuming the refractive index for air is $n_{\text{air}} \sim 1$, Snell’s law

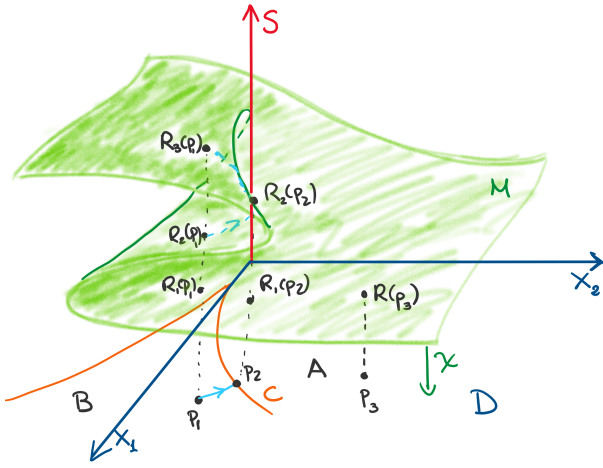


FIG. 2. The geometry of the cusp catastrophe. The generating function Φ is a quartic polynomial leading to a “stationary manifold” M . The solution space M is a cubic surface folding over itself in the space $S \times D \equiv \{a\} \times \{(x_1, x_2)\}$. The projection $\chi : M \rightarrow \{(x_1, x_2)\}$ becomes singular at points where a tangential direction to M is vertical (green curve in M). The image of these points under the projection is the caustic C . The caustic separates D in two regions: the interior of the caustic B where a point p_1 has three preimages in M and the exterior of the caustic A where a point p_3 has only preimage in M . This explains the higher luminosity in the interior of the caustic. On the caustic, points p_2 have two preimages, though one of them (resulting from the merger of two stationary points in M , $R_2(p_2)$ and $R_3(p_2)$ in the figure) is double.

relates the angles as $\sin i = n \sin r$ with n being the refractive index of water. The light travels inside the droplet and gets reflected in the internal side. Finally, the reflected ray reaches again the surface and is transmitted to the exterior with an angle i . Further inner reflections can be considered and lead to higher-orders “bows”. Denoting by D the deflected angle with the respect to the initial incident ray, it holds (cf. first panel of Fig. 3)

$$D(i) = 2i - 4r(i) + \pi. \quad (4)$$

In this problem, one can take as “state variable” \mathbf{R} the incident angle i though, for convenience one can take the “impact parameter” $b = a \sin i$. The natural “control parameter” \mathbf{X} is the angle D or, equivalently, the complementary angle $\theta = \pi - D$. Indeed, an observer will receive rays scattered for all droplets laying in a cone (centered in his or her eye and with axis pointing away from the Sun) of angle θ . To understand the rainbow phenomenon, we consider a ray incident perpendicularly at the droplet, that is $b = 0$. This ray is scattered back, i.e. $D = \pi$. As b increases the deflection angle D diminishes (see angles D of light rays A , B and C in the second panel of Fig. 3). The observer receives rays coming from all these directions. However, at a critical b_C the deflected angle reaches a minimum D^C , so that when b grows over b_C the angle D grows. This means that the observer does not receive any ray with $D < D^C$ (or equivalently, with $\theta > \theta_C$). Furthermore, the “density” of rays accumulated in this turning point D^C formally diverges, so the associated light intensity

at this angle also diverges, i.e. we find a caustic in the control space at D^C . The rainbow, for a fixed wavelength, is the arch seen by the observer corresponding to the light received from droplets along the cone of angle θ_C , for which an enhanced intensity occurs. Taking into account light dispersion in water, namely the fact that blue/violet colors are more scattered (larger n) than red colors, different bows forms for different angles, with $\theta_C^{\text{blue}} < \theta_C^{\text{red}}$, giving rise to the familiar rainbow structure (see right panel of Fig. 4).

From Eq. (4) one can derive an expression for the corresponding phase difference as [69]

$$\Phi_{\text{rainbow}}(\delta; \Delta)_{\text{lens}} = -\frac{1}{3}\alpha\beta\delta^3 + \beta\delta\Delta, \quad (5)$$

where $\delta = b - b_c$ and $\Delta = D - D_c$ are, respectively, the state variable and control parameter centered around the (critical) caustic values, the parameter α is $\alpha = \frac{1}{2}D''(b_c)$ and β an appropriate constant. The corresponding caustic is referred to as a ‘fold’ and it is illustrated in Fig. 5.

The rainbow is thus a caustic, namely the ‘fold caustic’, in the angle control parameter θ . If we consider now the issue of the illumination at both sides of the caustic, we observe that in the inner part of the rainbow ($\theta < \theta_C$) there are two scattered rays reaching the observer, whereas in the outer part ($\theta > \theta_C$) there is no scattered ray at all, so the exterior is actually dark. This is indeed observed in actual rainbows in the sky, where the inner part of the rainbow is more illuminated than the exterior. This is the same phenomenon we saw above in the (cusp) caustic of the lens, where couples of rays coalesce at the caustic (where the intensity diverges) and disappear on the other side. In the lens (cusp) case no full darkness is found in the exterior due to the persistence of a third ray.

Our BBH problem, with a passage from an oscillating signal to its total extinction after the merger (the ringdown is not accounted by this mechanism) seems mathematically more akin to the rainbow fold caustic rather than the lens cusp one. We will elaborate later on this point, once we identify the appropriate “control parameter” in our problem, corresponding formally to the deflecting angle θ .

B. The framework of Singularity theory

Let us introduce now systematically, though briefly, the relevant elements in the discussion within the framework of Singularity/Catastrophe theory. We shall aim to explain the abstraction which encompass the examples given above. Further details and references can be found in e.g. [68–74].

- i) *Generating function, stationary manifold and catastrophe map.* Let us denote by S a m -dimensional space of ‘state variables’ $\mathbf{R} = (R_1, \dots, R_m)$ and by D the n -dimensional space of ‘control’ variables $\mathbf{X} = (X_1, \dots, X_n)$. The a priori number of state variables can be large (possibly infinite, as in the case of an underlying field theory), but not all of those variables are relevant regarding the caustic structure. The number of independent state variables is reduced to the minimal

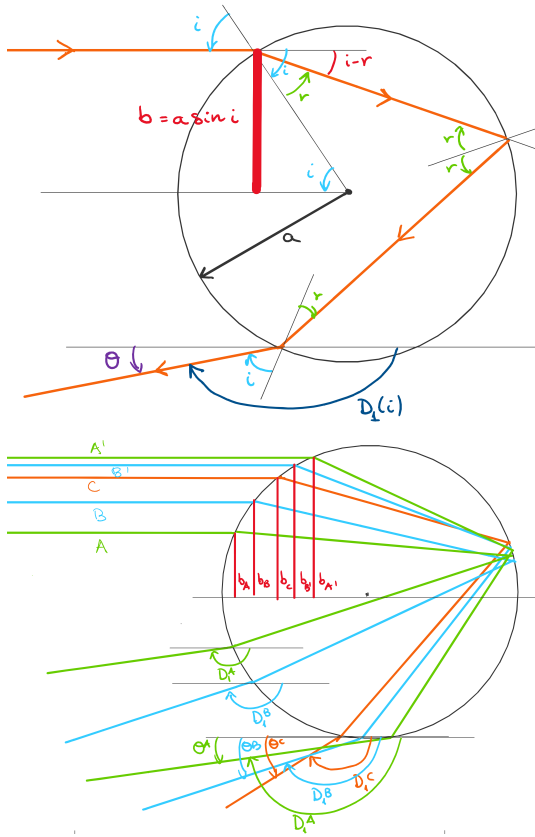


FIG. 3. Rainbow fold caustic: Parallel incident rays from the left are refracted and internally reflected by the water droplet. There is a critical value θ_C of the deflection angle (depending on the wavelength) which is responsible for the rainbow phenomenon.

number necessary to encode the (topological) structure of the catastrophe. We will focus here on situations with a finite, and in fact small, number m of state variables. The number of control parameters n could also be infinite [70], but we focus here on a finite number $n \geq m$. The important dimensionality is the dimension of the space of control parameters C minus the dimension of the singularity, known as the ‘‘co-dimension’’. For both the fold and cusp catastrophes the singularities are point-like (i.e. the fold and the tip of the cusp); thus the fold and cusp catastrophes have co-dimension 1 and 2 respectively, coinciding with the dimensions of the control space.

As in the examples presented above, the ray trajectories are obtained from the extremalisation of a ‘generating function’, $\Phi : S \times D \rightarrow \mathbb{R}$ with the structure:

$$\Phi(\mathbf{R}, \mathbf{X}) = \sum_{\alpha=1}^m R_{\alpha} X_{\alpha} + f(\mathbf{R}; X_{m+1}, \dots, X_n). \quad (6)$$

Extremizing with respect to the state variables

$$\nabla_{R_{\alpha}} \Phi = 0, \quad \alpha \in \{1, \dots, m\}, \quad (7)$$

permits us to write

$$X_{\alpha} = -\nabla_{R_{\alpha}} f, \quad \alpha \in \{1, \dots, m\}. \quad (8)$$

These m relations determine an n -dimensional manifold $M \subset S \times D$ of stationary points, the latter corresponding to actual rays reaching from the source to the detector. We will refer to M as the *stationary or catastrophe manifold* (cf. e.g. [74]). We can define a projection map $\chi : M \rightarrow D$ (‘catastrophe map’) as the restriction to M of the projection $\pi : S \times D \rightarrow D$

$$\pi(\mathbf{R}, \mathbf{X}) = \mathbf{X}, \quad \chi = \pi|_M. \quad (9)$$

We can locally parameterize $p \in M$ as $p = (R_1, \dots, R_m, X_{m+1}, \dots, X_n)$, by using Eq. (8), so

$$\begin{aligned} \chi(R_1, \dots, R_m, X_{m+1}, \dots, X_n) \\ = (X_1(\mathbf{R}), \dots, X_m(\mathbf{R}), X_{m+1}, \dots, X_n) \end{aligned} \quad (10)$$

- ii) *Magnification matrix and intensity.* If the map χ is one-to-one, for each chosen value of the control parameter \mathbf{X}_o it corresponds a unique ray with a given (source) state variable value \mathbf{R}_o and, in particular, a single-valued $\mathbf{R} = \mathbf{R}(\mathbf{X})$ exists. We can define the

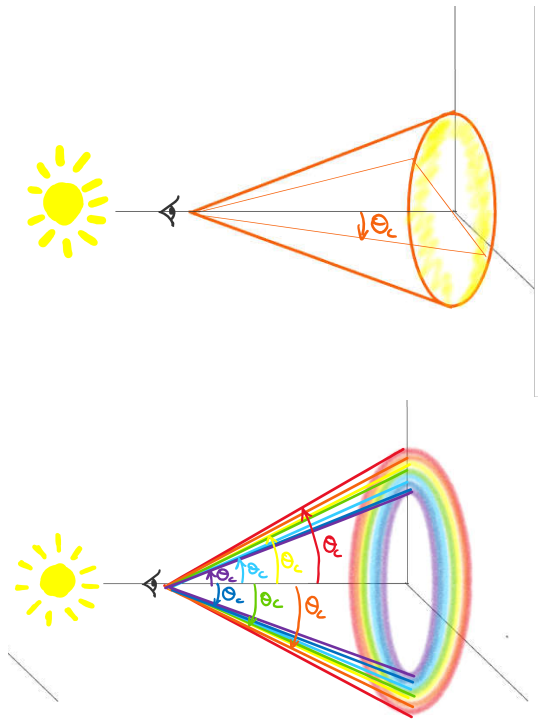


FIG. 4. Rainbow angular structure. The rainbow corresponds to a caustic at angle θ_C , where the density of scattered rays diverges. Only the interior of the cone with opening angle θ_C is illuminated (by two rays), whereas the exterior of the caustic remains dark. In the second panel, the dependence of the critical angle on the wavelength is taken into account, resulting in the familiar rainbow pattern.

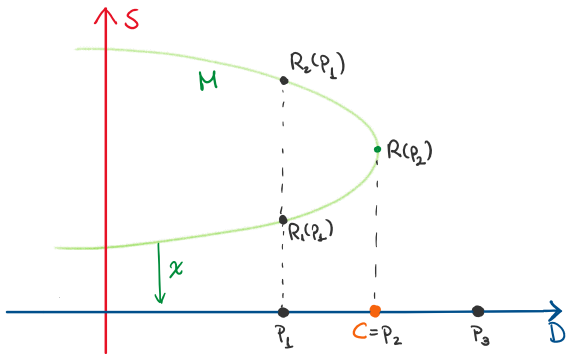


FIG. 5. Fold catastrophe: The generating function in this case is cubic, leading to a quadratic curve in the space (plane) $S \times D$. In this case the caustic C is just a point, the image under χ of the vertex of the parabola. “Interior” points p_1 to the caustic (to the left of C in the figure) have two preimages $R_1(p_1)$ and $R_2(p_1)$ in M , whereas “exterior” points p_2 have no preimage at all. Actually, points $R_1(p_1)$ and $R_2(p_1)$ in M merge at the caustic, disappearing on the other side: in the fold, we pass from illumination to complete absence of illumination at the caustic. The rainbow caustic is of this type [69], where the control variable X_1 is the deflection angle D (cf. Fig. 3).

magnification matrix $\mathcal{M}_{\alpha\beta}$, with $\alpha, \beta \in \{1, \dots, m\}$, of the rays from S to D as [68, 70]

$$\mathcal{M}_{\alpha\beta}(\mathbf{X}) = \left(\frac{\partial R_\alpha}{\partial X_\beta} \right)_{(X_{m+1}, \dots, X_n)} \quad (11)$$

$$= \left(\frac{\partial X_\alpha}{\partial R_\beta} \right)_{(X_{m+1}, \dots, X_n)}^{-1}(\mathbf{R}(\mathbf{X})). \quad (12)$$

Using Eqs. (6), (7) and (8) one can write

$$\mathcal{M}_{\alpha\beta}^{-1}(\mathbf{X}) = - \left(\frac{\partial^2 f}{\partial R_\alpha \partial R_\beta} \right)_{(X_{m+1}, \dots, X_n)}(\mathbf{R}(\mathbf{X})) \quad (13)$$

$$= - \left(\frac{\partial^2 \Phi}{\partial R_\alpha \partial R_\beta} \right)_{(X_1, \dots, X_n)}(\mathbf{R}(\mathbf{X})) \quad (14)$$

The intensity I of the light at the point \mathbf{X} in this (geometric optics) ray picture is then given by

$$I(\mathbf{X}) = |\det \mathcal{M}_{\alpha\beta}|. \quad (15)$$

- iii) ‘Branches’ of the stationary manifold M . In generic situations the manifold M folds over in $S \times D$ making the projection χ not injective, several ‘rays’ arrive at \mathbf{X} associated with different state configurations in the source, i.e. several state variables $\mathbf{R}_i \in S$ are in general associated with a given $\mathbf{X} \in D$ and the application $\mathbf{R}_i = \mathbf{R}_i(\mathbf{X})$ is multi-valued. In this situation, to get the total intensity we must correct expression (15) by summing over all rays reaching \mathbf{X} , so we must write

$$I(\mathbf{X}) = \sum_i |\det \mathcal{M}_{\alpha\beta}(\mathbf{R}_i(\mathbf{X}))|. \quad (16)$$

In this setting, we can consider the different branches S_i of an action function $S_i : D \rightarrow \mathbb{R}$ introduced as

$$S_i(\mathbf{X}) = \Phi(\mathbf{R}_i(\mathbf{X}), \mathbf{X}). \quad (17)$$

This defines the different branches of a (generically) multi-valued function solving a Hamilton-Jacobi equation [71] and whose constant-value surfaces constitute the geometric wavefronts in geometrical optics [75].

- iv) *Caustic*. The caustic C is precisely the set of points in D where the different branches (17) of a multi-valued ‘action function’ join. In other words, the caustic is given by the projection to D of the set of points in $S \times D$ where M “folds over itself”. Thus, two stationary points p_i and $p_j \in M$ that are separated on one side of the caustic, do coalesce $p_i \rightarrow p_j$ at the caustic and disappear on the other side.

Technically, the caustic C is the image by χ of the critical points of such projection χ , also referred to as ‘critical set of M ’ (namely the points in M where $d\chi$ is singular, i.e. where it has a non-trivial kernel). From the parametrization in (10) this happens when the Hessian of Φ (evaluated on M) is not invertible. Equivalently, from Eq. (13), this happens when the inverse of the magnification matrix $\mathcal{M}_{\alpha\beta}$ is singular and its determinant diverges. As a consequence of Eq. (15) the ‘light’ intensity I diverges at the caustic.

C. Universal diffraction pattern on caustics

As we have just seen, the geometric optics approximation to light propagation breaks down at the caustic, where the intensity diverges. Moreover, in our BBH setting we are fundamentally interested, apart from the intensity of the gravitational signal, also on the radiation wave field at the caustic. This requires to go beyond the geometric optics approximation. Without having to invoke the full Maxwell theory, a suitable solution to the problem is provided by Fresnel diffraction integrals [76]. Specifically, diffraction on caustics brings about two remarkable outcomes: i) it regularizes the intensity divergence, and ii) diffraction patterns show universal behavior only depending on the topological nature of the caustic.

The starting point is the Fraunhofer/Fresnel diffraction integral for the wave field $\psi(\mathbf{X})$ [70–72, 77, 78]

$$\psi(\mathbf{X}) = \left(\frac{k}{2\pi} \right)^{\frac{m}{2}} \int d^m \mathbf{R} e^{ik\Phi(\mathbf{R}, \mathbf{X})} a(\mathbf{R}, \mathbf{X}). \quad (18)$$

Here $k = 2\pi/\lambda$ with λ the wavelength, $\Phi(\mathbf{R}, \mathbf{X})$ is an appropriate generating function (6) accounting for the phase delay and $a(\mathbf{R}, \mathbf{X})$ is a slowly varying function in (\mathbf{R}, \mathbf{X}) corresponding to an attenuation factor along the ray path.

Let us consider first diffraction when we are away from a caustic, so stationary points p_i of $\Phi(\mathbf{R}, \mathbf{X})$ are not degenerate. In this situation, we can approximate Eq. (18) using a *stationary phase approximation*. The phase of the exponential in (18) oscillates rapidly and the main contribution to the

integral comes from the stationary points of Φ . Expanding up to second order around a given stationary point \mathbf{R}_i we write

$$\begin{aligned} \Phi(\mathbf{R}, \mathbf{X}) &= \Phi(\mathbf{R}_i(\mathbf{X}), \mathbf{X}) + \nabla_{R_\alpha} \Phi(\mathbf{R}, \mathbf{X})|_{\mathbf{R}_i(\mathbf{X})} R_\alpha \\ &+ \nabla_{R_\alpha} \nabla_{R_\beta} \Phi(\mathbf{R}, \mathbf{X})|_{\mathbf{R}_i(\mathbf{X})} R_\alpha R_\beta + O(\mathbf{R}^3) \end{aligned} \quad (19)$$

The linear term vanishes since we are at a stationary point. Neglecting terms beyond quadratic ones reduces (18) to a Gaussian integral contribution around each stationary point

$$\psi(\mathbf{X}) \sim \sum_i \frac{e^{i(kS_i(\mathbf{X}) + \alpha_i \pi/4)}}{|\det(\mathcal{M}_{\alpha\beta}^{-1}(\mathbf{X}))|^{\frac{1}{2}}} a(\mathbf{R}_i(\mathbf{X}), \mathbf{X}), \quad (20)$$

where α_i is the signature (i.e. the number of positive minus the number of negative eigenvalues, counting multiplicities) of the $\mathcal{M}_{\alpha\beta}^{-1}$ matrix, and $S_i(\mathbf{X}) := \Phi(\mathbf{R}_i(\mathbf{X}), \mathbf{X})$ is the action function in the branch corresponding to p_i . Taking into account that the field intensity satisfies $I \sim |\psi(\mathbf{X})|^2$, from this expression we read:

- i) For a single stationary point we recover (if ignoring the attenuation factor) the geometric optics value (15), namely $I \sim |\det(\mathcal{M}_{\alpha\beta})|$.
- ii) When multiple non-degenerate stationary points exist, the geometric optics expression (16) is corrected with additional interference terms controlled by the the action function evaluated on the stationary points.
- iii) At the caustic the intensity diverges, since $\det(\mathcal{M}_{\alpha\beta}^{-1}(\mathbf{X})) = 0$. The stationary phase approximation is not valid here, and it is necessary to go to higher orders in the expansion (19) and/or modify the diffraction treatment.

1. Diffraction on caustics: ‘uniform asymptotic approximation’

As we have just seen, when stationary points p_i merge at the caustic the direct stationary phase approximation is no longer valid and must be modified. A ‘uniform asymptotic approximation’, valid both for isolated and merged extrema has been systematically developed for different caustics starting with the work of Chester, Friedman and Ursell on the fold catastrophe [79], later extended to higher caustics (see references in [70–72]). The subject was put on a rigorous basis by Duistermaat [77] (see also [80, 81]). The method depends critically on the topological features of the process of merging of stationary points p_i in M . The remarkable fact is that, for finite co-dimension n , the *structurally stable* topological possibilities are classified by the celebrated Arnol’d-Thom theorem in singularity theory [82, 83]. Here “structurally stable” refers to stability under generic small perturbations. In the context of caustics in optical theory, that we shall discuss further shortly, the concept of structural stability represents effects such as deformations of water droplets due to gravity, propagation effects in the surrounding medium etc. These effects make the physical situation non-ideal but do not change the essential topological features of the caustic itself.

As an application, this theorem classifies structurally stable caustics in terms of the generating function $\Phi(\mathbf{R}, \mathbf{X})$, that can be reduced through appropriate re-definitions of state variables and control parameters to standard universal forms $\phi(\mathbf{r}, \mathbf{x})$. Table I presents all the possibilities for co-dimension $n \leq 3$. Each of these standard universal forms is a polynomial in the state and control variables; it consists of a “germ”, i.e. a polynomial in the state variable only which contains the most singular part. The rest of the polynomial is linear in the control variables and describes the “unfolding” of the singularity away from the germ. In this setting, the above-commented ‘uniform asymptotic approximation’ relies on the comparison of the integral expression (18) for $\psi(\mathbf{X})$ with an integral

$$j(\mathbf{x}) = \left(\frac{k}{2\pi}\right)^{\frac{m}{2}} \int d^m \mathbf{r} e^{ik\phi(\mathbf{r}, \mathbf{x})}, \quad (21)$$

defined in terms of generating functions $\phi(\mathbf{r}, \mathbf{x})$ that share with $\Phi(\mathbf{R}, \mathbf{X})$ the topological structure of the process of “coalescence of stationary points”, but such that the choice of state variables \mathbf{r} and control parameters \mathbf{x} make the functional form of $\phi(\mathbf{r}, \mathbf{x})$ simpler than that of $\Phi(\mathbf{R}, \mathbf{X})$. Under the light of the Arnol’d-Thom theorem, a natural choice for $\phi(\mathbf{r}, \mathbf{x})$ is given by those ones corresponding to “elementary catastrophes” in Table I.

As a consequence, the Arnol’d-Thom theorem not only classifies elementary caustics in geometric optics, but also the universal diffraction patterns on caustics that “clothe the skeleton provided by stable caustics” [70] when the wave nature of light is taken into account. In our present setting, if caustics provide the ‘skeleton’ of the BBH merger waveform model, diffraction can be paraphrased as providing its ‘flesh’.

Let us sketch the procedure to express (18) in terms of (21) by closely following [70] (we bring attention to the recent work [84] for the application of these uniform asymptotic approximations in a gravitational setting [85]). We start by considering a generating function Φ encoding a singularity structure “equivalent” to one of the “elementary catastrophes”, i.e. equivalent to one of the generating functions ϕ in Table I. More precisely, Φ and ϕ are equivalent in the sense (cf. e.g. [74]) that there exist i) for each fixed \mathbf{X} , a diffeomorphism $\mathbf{R} \mapsto \mathbf{r}(\mathbf{R})$ from the natural state variables in the problem \mathbf{R} to the standard ones \mathbf{r} , ii) a diffeomorphism between control parameters $\mathbf{X} \mapsto \mathbf{x}(\mathbf{X})$, and iii) a smooth function $C(\mathbf{X})$ (the “shear term”) such that

$$\Phi(\mathbf{R}, \mathbf{X}) = C(\mathbf{X}) + \phi(\mathbf{r}(\mathbf{R}), \mathbf{x}(\mathbf{X})). \quad (22)$$

The function C must be, in particular, smooth at the caustic. To fix $C(\mathbf{X})$ and $\mathbf{x}(\mathbf{X})$, we can use that critical points of Φ are mapped onto critical points of ϕ , which follows from the fact that C is independent of the state variables. Then, since for co-dimension n there can be $n + 1$ critical points involved in the coalescence [70], we can impose

$$\Phi(\mathbf{R}_i(\mathbf{X}), \mathbf{X}) = C(\mathbf{X}) + \phi(\mathbf{r}_i(\mathbf{x}(\mathbf{X})), \mathbf{x}(\mathbf{X})) \quad i \in \{1, \dots, n+1\}. \quad (23)$$

For a fixed \mathbf{X} this provides $n + 1$ conditions to determine the $n + 1$ numbers $C(\mathbf{X})$ and $\mathbf{x}(\mathbf{X}) = (x_1(\mathbf{X}), \dots, x_n(\mathbf{X}))$.

Co-dimension	Caustic name	Generating function $\phi(\mathbf{r}, \mathbf{x})$
1	Fold	$r_1 x_1 + r_1^3$
2	Cusp	$r_1 x_1 + \frac{1}{2} x_2 r_1^2 + \frac{1}{4} r_1^4$
3	Swallow tail	$r_1 x_1 + \frac{1}{2} x_2 r_1^2 + \frac{1}{3} x_3 r_1^3 + \frac{1}{5} r_1^5$
3	Elliptic umbilic	$-r_1 x_1 - r_2 x_2 - x_3 (r_1^2 + r_2^2) - 3r_1 r_2^2 + r_1^3$
3	Hyperbolic umbilic	$-r_1 x_1 - r_2 x_2 + x_3 r_1 r_2 + r_1^3 + r_2^3$

TABLE I. “Elementary catastrophes” classifying, in particular, structurally stable caustics with co-dimension $n \leq 3$. Generating functions $\phi(\mathbf{r}, \mathbf{x})$ share the form of $\Phi(\mathbf{R}, \mathbf{X})$ in (6), with state variables $\mathbf{r} = (r_1, r_2, \dots, r_m)$ and control parameters $\mathbf{x} = (x_1, x_2, \dots, x_n)$, with $m \leq n$, but lower-case letters are chosen to emphasize the fact that these are standard universal forms to which equivalent $\Phi(\mathbf{R}, \mathbf{X})$ can be reduced under appropriate diffeomorphisms.

Inserting Eq. (22) into the diffraction integral expression (18) for $\psi(\mathbf{X})$, we can write

$$\psi(\mathbf{X}) = \left(\frac{k}{2\pi} \right)^{\frac{m}{2}} e^{ikC(\mathbf{X})} \int d^m \mathbf{r} g(\mathbf{r}, \mathbf{X}) e^{ik\phi(\mathbf{r}, \mathbf{x}(\mathbf{X}))}, \quad (24)$$

where

$$g(\mathbf{r}, \mathbf{X}) = \left| \frac{d\mathbf{R}}{d\mathbf{r}} \right|(\mathbf{X}) a(\mathbf{R}(\mathbf{r}), \mathbf{X}), \quad (25)$$

with $\left| \frac{d\mathbf{R}}{d\mathbf{r}} \right|(\mathbf{X})$ the Jacobian of the map $\mathbf{r} \mapsto \mathbf{R}(\mathbf{r})$, for fixed \mathbf{X} .

Then the function $g(\mathbf{r}, \mathbf{X})$ is approximated by exploiting the fact that the main contribution comes from the stationary points (since $k \gg 1$). Following [70], we first write

$$g(\mathbf{r}, \mathbf{X}) = g_0(\mathbf{x}(\mathbf{X})) + g_1(\mathbf{x}(\mathbf{X})) \cdot \nabla_{\mathbf{x}} \phi(\mathbf{r}, \mathbf{x}) + h(\mathbf{r}(\mathbf{x}(\mathbf{X}))) \cdot \nabla_{\mathbf{r}} \phi(\mathbf{r}, \mathbf{x}), \quad (26)$$

with $g_1(\mathbf{x}(\mathbf{X}))$ and $h(\mathbf{r}(\mathbf{x}(\mathbf{X})))$ appropriate n -dimensional and m -dimensional vectors, respectively. The approximation then consists in estimating the third term by its value at the critical points [86]. Using the relation (23) this last term vanishes and we can write

$$g(\mathbf{r}, \mathbf{X}) \sim g_0(\mathbf{x}(\mathbf{X})) + g_1(\mathbf{x}(\mathbf{X})) \cdot \nabla_{\mathbf{x}} \phi(\mathbf{r}, \mathbf{x}). \quad (27)$$

Finally, inserting (27) into (24) we can express $\psi(\mathbf{X})$ in terms of diffraction caustics integrals $j(\mathbf{x})$ defined in Eq. (21) and their derivatives

$$\psi(\mathbf{X}) = e^{ikC(\mathbf{X})} \left(g_0(\mathbf{x}(\mathbf{X})) j(\mathbf{x}(\mathbf{X})) + \frac{g_1(\mathbf{x}(\mathbf{X}))}{ik} \cdot \nabla_{\mathbf{x}} j(\mathbf{x}(\mathbf{X})) \right). \quad (28)$$

For explicitly known functions $\Phi(\mathbf{R}, \mathbf{X})$ and $a(\mathbf{R}, \mathbf{X})$ a further step can be taken [70] to determine, for each \mathbf{X} , the $n+1$ numbers $g_0(\mathbf{x}(\mathbf{X}))$ and $g_1(\mathbf{x}(\mathbf{X}))$ [87]. Although of genuine interest, we will not need this further step at the present stage of the discussion of the BBH merger waveform model.

We would like to conclude this section with some remarks:

i) *Universal diffraction pattern.* Expression (28), captures the functional form of the wave field in terms of universal diffraction patterns on elementary caustics (even in the case when functions $\Phi(\mathbf{R}, \mathbf{X})$ and $a(\mathbf{R}, \mathbf{X})$ are not known). This expression, consequence of the ‘uniform asymptotic approximation’, is the main result we want to focus on here, as the key ingredient for a template proposal for BBH merger waveforms.

ii) *Transitional approximation.* The full uniform asymptotic approximation is valid on and near the caustic, but also far from it. If we focus on the region close to the highest-order singularity in the caustic, the second term in (28), involving the derivative of the elementary diffraction, can be neglected [70] (also \mathbf{x} could be taken as $\mathbf{x}(\mathbf{X}) = \mathbf{X}$). This defines the so-called *transitional approximation* [88].

On the other hand, if we are interested in connecting with the behavior far from the caustic, a stationary phase expansion on the critical points of ϕ leads to a recovery of the expression (20) (cf. e.g. [70–72, 89]).

iii) *Diffraction catastrophe universal scaling laws.* As $k \rightarrow \infty$, the diffraction pattern must recover the geometric optics limit, in particular the divergence at the caustic and the vanishing of the spacing in the diffraction fringe patterns. This is indeed the case and, moreover, such behavior is universal and controlled by scaling laws only depending on the caustic type [39, 71, 72, 90]. Specifically, introducing the *elementary diffraction catastrophe*

$$J(\mathbf{x}) = \left(\frac{1}{2\pi} \right)^{\frac{m}{2}} \int d^m \mathbf{r} e^{i\phi(\mathbf{r}, \mathbf{x})}, \quad (29)$$

not depending on k , the asymptotic solution for large k satisfies the scaling (self-similar) law [91]

$$j(x) = k^\beta J(k^{\sigma_i} x_i), \quad (30)$$

with scaling exponents β and σ_i , with $i \in \{1, \dots, n\}$ (n being the co-dimension of the caustic). The exponent β , referred to as *singularity index* and introduced

by Arnol'd, controls the divergence of the wave amplitude close to the highest order singularity of the caustic

$$|\psi(\mathbf{X})| \sim k^\beta. \quad (31)$$

On the other hand, the exponents σ_i (introduced by Berry [90]) control the fringe spaces along the control variables x_i as k diverges. In particular, the so-called ‘fringe index’ γ is defined as $\gamma = \sum_i^n \sigma_i$ (cf. [90]).

- iv) *Relevant and irrelevant state variables.* Given a control space of dimension n , the number of state variables m is constrained by $m \leq n$ (cf. e.g. table I). Additional state variables do not change the diffraction catastrophe analysis, since they enter quadratically in the generating function Φ and therefore do not contribute at the caustic. In other words, the physical system could be described by a large (infinite, in a field theory) number of state variables and still the relevant state variables remain finite and constrained by the caustic codimension.

III. FOLD CAUSTICS AND BINARY BLACK HOLE MERGERS

A. Assumptions in the BBH caustic model

We are now in a position to formulate the elements of a model for the BBH merger waveform, based on the diffraction of gravitational waves on a caustic. The model is built on the following assumptions:

1. *Main Ansatz.* BBH gravitational wave propagation from the source to the detector can be modeled in a geometric-optics approximation in terms of the extremalization of a phase-difference function $\Phi(\mathbf{R}, \mathbf{X})$, with control parameters given by the spacetime coordinates of the detector. The transition from the growing intensity of the signal in the inspiral phase to the disappearance of the signal after the merger can be described in terms of a caustic (catastrophe).
2. *Co-dimension $n = 1$: fold caustic.* Along the world-line of the detector, the only relevant control parameter is the time, so that $\mathbf{X} = t$. According to Ansatz 1 and the Arnol'd-Thom theorem, the catastrophe should be a fold-caustic and, as a consequence: i) there is a single (relevant) state variable R for the whole source system, and ii) the phase difference $\Phi(R, t)$ is topologically equivalent to $\phi(r, \tau) = \frac{1}{3}r^3 + \tau r$ with r and τ obtained by reparametrizations of (R, t) .
3. The natural range of the parameter r is assumed to be to full real line, i.e. $r \in (-\infty, \infty)$. No other assumptions are made on R or r at this stage, which remain as abstract ‘effective source variables’ [92]. Spacetime point labels (namely t here) are ‘control parameters’, and not ‘state variables’, in a spirit akin with the general relativistic perspective.

B. Diffraction on a fold caustic: the Airy uniform approximation

Under the previous assumptions, we model each of the two polarizations of the gravitational waveform as the Fraunhofer/Fresnel diffraction regularization of the fold-caustic geometric optics description of the merger (cf. the treatment in [72] for the electromagnetic case). Specifically, we replace Φ with its topologically equivalent ϕ , and begin with the integral

$$j(\tau) = \left(\frac{\omega}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} dr e^{i\omega(\frac{1}{3}r^3 + \tau r)}. \quad (32)$$

Notice that, for notational reasons, we have substituted the ‘wave number’ k parameter in Eq. (21) by the ‘frequency’ parameter ω , to stress that the caustic happens in time, and not in space. The diffraction integral $j(\tau)$ in Eq. (32) is directly related to the Airy function, which is defined as

$$\text{Ai}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(\tau r + r^3/3)} dr, \quad (33)$$

(notice the relation $\text{Ai}(\tau) = (1/\sqrt{2\pi})J(\tau)$, with $J(\tau)$ the universal caustic diffraction pattern in Eq. (29), for the particular fold case). On the one hand, making the change $\omega r^3 = \tilde{r}^3$ we can rewrite the ω -dependent $j(\tau)$ in terms of $\text{Ai}(\tau)$ as

$$\begin{aligned} j(\tau) &= \omega^{\frac{1}{6}} \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} d\tilde{r} e^{i(\frac{1}{3}\tilde{r}^3 + \omega^{\frac{2}{3}}\tau\tilde{r})} \\ &= \omega^{\frac{1}{6}} (2\pi)^{\frac{1}{2}} \text{Ai}(\omega^{\frac{2}{3}}\tau). \end{aligned} \quad (34)$$

On the other hand, considering the BBH phase difference $\Phi(R, t)$ in Ansatz 2, with appropriate diffeomorphisms $r(R)$ and $\tau(t)$ rendering it topologically equivalent to the canonical fold-generating function $\phi(r, \tau)$ in Table I, we can apply the analysis in section II C to write each gravitational wave polarization $\psi(t)$ in terms of Ai and Ai' . Specifically, applying Eq. (28) to the fold-caustic case, we get

$$\begin{aligned} \psi(t) &= e^{i\omega C(t)} \left(\omega^{\frac{1}{6}} \tilde{g}_0(t) \text{Ai}(\omega^{\frac{2}{3}}\tau(t)) \right. \\ &\quad \left. - i\omega^{-\frac{1}{6}} \tilde{g}_1(t) \text{Ai}'(\omega^{\frac{2}{3}}\tau(t)) \right). \end{aligned} \quad (35)$$

Two remarks are in order:

- i) *Fold-diffraction critical exponents.* Comparing Eq. (34) with the self-similar scalings in Eq. (30) we can identify the Arnol'd parameter singularity index β and the fringe index $\gamma = \sigma_\tau$ for the fold-diffraction, respectively $\beta = \frac{1}{6}$ and $\gamma = \frac{2}{3}$. In particular, the singularity index β predicts that, in the vicinity of the merger, the amplitude $|\psi|$ that provides the envelope of the gravitational wave signal presents the following power-law dependence in the frequency ω

$$|\psi(\omega)| \sim \omega^{\frac{1}{6}}. \quad (36)$$
- iii) *Airy pattern beyond ‘catastrophe optics’.* A tenet in asymptotic reasoning [39, 93] is that asymptotic limits (in the present case, geometric optics limit $\omega \rightarrow$

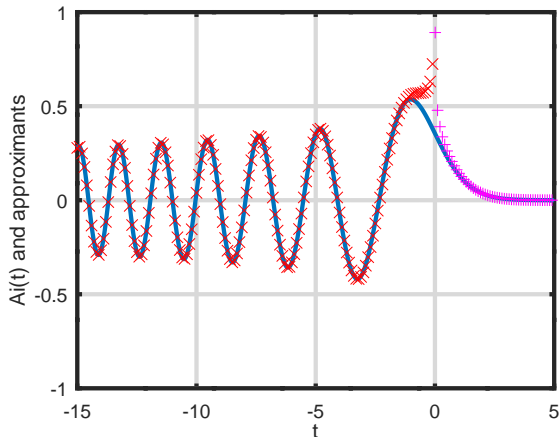


FIG. 6. The behavior of the Airy function on the real axis. Also shown are the two approximations from Eq. (39) for negative times (red) and from Eq. (40) for positive times (magenta).

∞) permit to identify the relevant pattern underlying the studied physical mechanism, the pattern itself remaining often valid away of that limit. This does not mean that the identified pattern accounts for the detailed quantitative aspects, but rather that it captures the relevant qualitative features. In this spirit, we promote expression (35) to the following ω -independent pattern

$$\psi(t) = e^{iC(t)} (g_0(t) \text{Ai}(\tau(t)) - i g_1(t) \text{Ai}'(\tau(t))) . \quad (37)$$

This ω -independent pattern will be indeed recovered in a completely independent approach in [65]. For our purposes here, this will be the starting point to be further development below which will lead to our model for the gravitational wave signal.

1. Some properties of the Airy function

We now enumerate some properties of the Airy function that we shall use below. We start with the Airy function defined in Eq. (33) for real arguments and satisfying the equation $\text{Ai}''(t) = t \text{Ai}(t)$. The Airy function can be extended as a holomorphic function $\text{Ai}(z)$ over the entire complex plane satisfying the Equation

$$\frac{d^2 \text{Ai}(z)}{dz^2} - z \text{Ai}(z) = 0 . \quad (38)$$

From a steepest descent evaluation of the Airy function, we have the following useful asymptotic expansions [94]. For (real) t large and negative, it holds

$$\text{Ai}(t) \sim \frac{1}{\sqrt{\pi}} \frac{1}{(-t)^{1/4}} \cos \left(\frac{2}{3} (-t)^{3/2} - \frac{\pi}{4} \right) . \quad (39)$$

We notice the factor $(-t)^{-1/4}$ controlling the asymptotic growth of the Airy function in this ‘early’ asymptotic

limit. In particular, although the waveform phase requires a reparametrization which is not determined so far, the familiar amplitude growth as $(-t)^{1/4}$ in the BBH inspiral phase is exactly captured by Airy asymptotics. As a matter of fact, already the interference terms in Eq. (20) account for this behavior through the determinant of the matrix $\mathcal{M}_{\alpha\beta}$. This is in a tantalizing agreement with the BBH inspiral behavior arising as a result of combining energy balance and Kepler’s laws and, more generally, post-Newtonian calculations.

For positive and sufficiently large t , we have

$$\text{Ai}(t) \sim \frac{1}{2\sqrt{\pi}} \frac{1}{t^{1/4}} e^{-\frac{2}{3} t^{3/2}} . \quad (40)$$

Thus, the decay is super-exponential for positive t . A plot of the Airy function on the real axis and these two asymptotic approximations is shown in Fig. 6. These asymptotic expansions yield excellent approximations to the Airy functions, and diverge from the true values only for a small interval around $t = 0$. In fact, Eq. (39) is seen to provide an excellent approximation even up to the peak of the Airy function, though it fails to capture the subsequent decay. The first peak of the Airy function, i.e. the first zero of $\text{Ai}'(t)$, occurs at $t_1 \approx -1.019$ and the peak value of the Airy function is $\text{Ai}(t_1) \approx 0.5357$. The value of the Airy function at $t = 0$ is $\text{Ai}(0) \approx 0.3550$ (we note that an inflection point of the Airy function occurs at $t = 0$, a feature of potential interest for later signal modeling).

It will be important in the Airy model below to be able to apply a phase shift to the Airy function. To this end, we define the functions $\text{Ai}_+(z)$ and $\text{Ai}_-(z)$ as follows

$$\text{Ai}_+(z) = \text{Ai} \left(e^{-2\pi i/3} z \right) , \quad \text{Ai}_-(z) = \text{Ai} \left(e^{2\pi i/3} z \right) . \quad (41)$$

It can then be shown that (see e.g. [94, 95])

$$\text{Ai}(z) = e^{\pi i/3} \text{Ai}_+(z) + e^{-\pi i/3} \text{Ai}_-(z) . \quad (42)$$

It is also easy to see that $\overline{\text{Ai}_+(\bar{z})} = \text{Ai}_-(z)$. Thus, along the real axis, where we write $z = t$, we will have $\text{Ai}_+(t) = \text{Ai}_-(t)$ and

$$\begin{aligned} \text{Ai}(t) &= e^{\pi i/3} \text{Ai}_+(t) + e^{-\pi i/3} \text{Ai}_-(t) \\ &= 2\Re \left[e^{\pi i/3} \text{Ai}_+(t) \right] . \end{aligned} \quad (43)$$

We can obtain steepest descent approximations to $A_{\pm}(z)$. For example, in the region in the complex plane with $|\text{Arg}(z) - 2\pi/3| < (1 - \epsilon)\pi$ (for some $\epsilon > 0$), we have

$$\text{Ai}_+(z) \sim \frac{e^{\pi i/6}}{2\sqrt{\pi} z^{1/4}} e^{-\frac{2i}{3} (-z)^{3/2}} . \quad (44)$$

Since this is of the form of a complex exponential, the form of the Airy function given in Eq. (43) allows us to define a phase shift of $\pi/2$ in the Airy function (just as the sine and cosine functions have a phase offset of $\pi/2$)

$$\begin{aligned} \text{Ai}_{-\pi/2}(t) &:= e^{\pi i/3} e^{-i\pi/2} \text{Ai}_+(t) + e^{-\pi i/3} e^{i\pi/2} \text{Ai}_-(t) \\ &= 2\Im \left[e^{\pi i/3} \text{Ai}_+(t) \right] . \end{aligned} \quad (45)$$

The following asymptotic expansions for large negative t then follows

$$\text{Ai}_{-\pi/2}(t) \sim \frac{1}{\sqrt{\pi}} \frac{1}{(-t)^{1/4}} \sin\left(\frac{2}{3}(-t)^{3/2} - \frac{\pi}{4}\right). \quad (46)$$

Comparing with the asymptotic expansion for $\text{Ai}(t)$ given in Eq. (39), we see explicitly the phase shift of $\pi/2$. It is straightforward to introduce a general phase shift φ_0

$$\text{Ai}_{\varphi_0}(t) := e^{\pi i/3} e^{i\varphi_0} \text{Ai}_+(t) + e^{-\pi i/3} e^{-i\varphi_0} \text{Ai}_-(t). \quad (47)$$

The phase shifted functions $\text{Ai}_{\varphi_0}(t)$ and $\text{Ai}_{\varphi_0-\pi/2}(t)$ shall be used below to model the two polarizations of the gravitational wave signal.

IV. THE AIRY MODEL FOR COMPACT BINARY MERGERS

A. Transitional approximation to Airy-diffraction binary merger waveforms

In the previous sections we have argued that the gravitational wave signal for a generic compact binary merger (excluding the post-merger phase) should be described in terms of the (amplitude) modulation and reparameterization of the Airy function and its derivative, as captured in Eq. (37). Conventionally, the two polarizations of the gravitational wave signal from a coalescing binary are written in terms of reparameterized sine and cosine functions

$$h_+(t) = A_+ \eta(t) \cos \varphi(t), \quad (48)$$

$$h_\times(t) = A_\times \eta(t) \sin \varphi(t). \quad (49)$$

Here $\eta(t)$ is a slowly varying function which determines the variation of the amplitude, while the phase $\varphi(t)$ is a rapidly varying function of t . For the dominant quadrupole mode ($\ell = m = 2$), the amplitudes $A_{+, \times}$ are

$$A_+ = \frac{1 + \cos^2 \iota}{2}, \quad A_\times = \cos \iota, \quad (50)$$

where ι is the inclination of the angular momentum of the binary with the line-of-sight to the detector. The parameters of the binary system appear in $\eta(t)$ and $\varphi(t)$. Let m_1 and m_2 be the masses of the two black holes, $M = m_1 + m_2$ the total mass, $\mu = m_1 m_2 / M$ the reduced mass, and $\mathcal{M} = \mu^{3/5} M^{2/5}$ the chirp mass. Let D be the distance to the source. Then, within the post-Newtonian approximation we have

$$\eta(t) = \frac{GM}{c^2 D} \left(\frac{t_c - t}{5GM/c^3} \right)^{-1/4}. \quad (51)$$

At the coalescence time t_c , the separation between m_1 and m_2 vanishes, and the amplitude diverges. The phase $\varphi(t)$ depends on the masses and also on the spins of the two black holes. At leading order

$$\varphi(t) = \varphi_0 - 2 \left(\frac{t_c - t}{5GM/c^3} \right)^{5/8}. \quad (52)$$

The coalescence phase φ_0 is the phase when $t = t_c$. Higher order corrections including also other parameters such as spin and eccentricity are well known in the literature.

As we approach the coalescence time, $\eta(t)$ varies increasingly rapidly and eventually diverges. Thus, the distinction between the amplitude and phase as being respectively slowly and rapidly varying is no longer valid. As discussed earlier, we postulate the underlying mechanism of this divergence to be akin to the divergence of the light intensity at a caustic in the geometric optics approximation.

Summarizing the discussion from sections II and III, to regularize this divergence we thus need to invoke the relevant Fresnel integral with an appropriate choice of phase function Φ ‘topologically equivalent’ to one of the canonical phase functions ϕ taken from Table I. In our case, the relevant control parameter is just the time t at which the detector collects data. Since we just have then a single control parameter it follows from our hypothesis and the Arnol’d-Thom theorem that the function Φ must be equivalent to the ϕ generating function for a fold catastrophe, which eventually leads an expression in terms of the Airy function and its derivative. We start then with the result of Eq. (37), namely a *uniform* approximation to the diffraction integral.

If we focus on the merger waveform, it is appropriate to consider the transitional approximation discussed in point ii) of section II C, namely we drop the second term in Eq. (37), involving the derivative of the Airy function [70, 88]. Remarkably, in this particular case involving a one-dimensional control space, it can be justified from the matching of the uniform approximation with the early (interference of ‘ray fields’) expression (20) valid far away from the caustic, that the second term in derivatives of the Airy function vanishes exactly [72, 89]. This permits to extend the transitional approximation for all the signal and work with a modulated and reparametrized Airy function.

B. The Airy model for binary mergers

In practical terms, by dropping the Ai' term in Eq. (37) through the transitional approximation, the role of the Airy function becomes that of replacing the sine and cosine in Eqs. (48) and (49). This regularizes the divergence of $\eta(t)$ at the coalescence time and accounts for the fact that we may not have a clean separation between the conventional amplitude and phase at the merger appearing in Eqs. (48) and (49). Since we are allowed to perform appropriate diffeomorphisms of the control parameters, we are led to reparameterizations of the phase shifted Airy functions $\text{Ai}_{\varphi_0}(\tau(t))$ and $\text{Ai}_{\varphi_0-\pi/2}(\tau(t))$ defined above, respectively the analogues of the cosine and sine functions in Eqs. (48) and (49). In this setting, we propose to modify Eqs. (48) and (49) as

$$h_+(t) = A_+ a(t) \text{Ai}_{\varphi_0}(\tau(t)), \quad (53)$$

$$h_\times(t) = A_\times a(t) \text{Ai}_{\varphi_0-\pi/2}(\tau(t)). \quad (54)$$

This is the actual model we propose for the BBH waveform merger. The amplitude function $a(t)$ appearing here is distinct from $\eta(t)$ and remains finite at the merger. For dimen-

sional reasons, it must still be proportional to $\frac{GM}{c^2 D}$, but the dependence on t is modified. This ansatz serves to describe a general elliptically polarized gravitational wave. In the conventional model of Eqs. (48) and (49), the binary parameters enter the phase $\varphi(t)$. Similarly, the system parameters enter the Airy function model via the reparameterization function $\tau(t)$ [96]. Note also that this model replaces the conventional complex signal $\eta(t)e^{i\varphi_0}e^{i\varphi(t)}$ with $e^{\pi i/3}e^{i\varphi_0}a(t)Ai_+(t)$. In regimes where the amplitude changes much more slowly than the phase, both representations are equivalent and in fact the conventional representation is more convenient. At the merger however the situation is different and the Airy description becomes necessary.

1. Early Airy asymptotics and BBH inspiral

At early times in the inspiral, it is not difficult to relate the early asymptotics of the Airy function to the standard post-Newtonian expressions. To this end, compare the argument of the \cos function in Eq. (39) with the post-Newtonian phase given in Eq. (52). For large $-t$ (so that both the t_c and constant terms can be ignored), we obtain

$$(-\tau)^{3/2} = 3 \left(\frac{-t}{5GM/c^3} \right)^{5/8}, \quad (55)$$

which implies

$$-\tau = 3^{2/3} \left(\frac{-t}{5GM/c^3} \right)^{5/12}. \quad (56)$$

The amplitude is then also easily obtained

$$a(t) \propto \left(\frac{-\tau(t)}{-t} \right)^{1/4} \propto (-t)^{-7/48}. \quad (57)$$

We see that $a(t) \rightarrow 0$ as $t \rightarrow -\infty$.

2. The merger time as the caustic time $\tau = 0$

Here we go beyond the post-Newtonian approximation and discuss the merger itself. It is conventional in the gravitational wave literature to denote the peak of the waveform as the merger. However, if one accepts the point of view proposed here, this may be misleading. The peak of the Airy function occurs at (dimensionless) $\tau \approx -1.019$ and the merger, in the underlying caustic model, a little later at the ‘‘caustic time’’ $\tau = 0$. The latter is the time at which the amplitude formally diverges in the ‘‘geometric optics’’ description and when the transition from oscillatory to damped behavior occurs. We shall adopt this perspective and shall use the terms ‘‘merger time’’ or ‘‘caustic time’’ for $\tau = 0$ [97].

Since the most accurate calculation of the merger signal to date is via numerical simulations, we can compare the Airy function with a numerical relativity signal. We take a particular signal from the SXS catalog [4] corresponding to the

merger of two non-spinning black holes. This is a ‘‘hybrid’’ waveform, constructed by stitching together a numerical relativity waveform with the corresponding post-Newtonian signal. We align the peak of the waveform with the largest peak of the Airy function. Matching then all preceding peaks and troughs of the Airy function for increasingly negative times with the corresponding peaks and troughs of the $+$ polarization of the Numerical waveform leads to the mapping $\tau(t)$ shown in the second panel of Fig. 7. Since the Airy function has no oscillations or zero-crossings at later times after the peak, this procedure cannot be extended to later times, and we must necessarily resort to extrapolation. We see that the function $\tau(t)$ obtained from this procedure increases rapidly as $\tau \rightarrow 0$. This is a consequence of the fact that the frequency of the Airy function decreases to zero as $\sqrt{\tau}$ while the frequency of the NR waveform increases monotonically. Similarly, the amplitude also increases rapidly, as shown in the third panel of Fig. 7. Nevertheless, the amplitude is a ratio of finite quantities, and it remains perfectly finite.

A remarkable fact of curves $\tau(t)$ and $a(t)$ in, respectively, the second and third panels of Fig. 7 is that the values of τ and a remain finite at the caustic time t_c (corresponding to $\tau = 0$), apparently reaching that value with an infinite slope. This suggests that such functions $\tau(t)$ and $a(t)$ are actually the lower branch of a multivalued function [98]. This would mean that the mapping $t \mapsto \tau$ does not extend for $t > t_c$ (namely $\tau = 0$), this entailing the impossibility of realizing the Airy diffraction BBH model beyond the caustic/merger point, consistently with our initial set up of the problem, where the ringdown must be accounted by a different physical mechanism and suggests strongly that the ringdown mechanism and phase must be matched at the merger/caustic time $\tau = 0$.

3. Transition to the ringdown at the caustic time

Even though the Airy function model aims mainly at the merger phase, the uniform approximation makes it also valuable in the inspiral phase. On the contrary, as we have just mentioned, the late ringdown phase is, as a matter of principle, not a part of this model. It requires an entirely different physical mechanism incorporating linear scattering resonances (quasi-normal modes) of a finite size resonator (the final merged black hole), in a linear ringdown phase. Nevertheless, assuming the validity of the Airy function model, we can still draw some conclusions about the ringdown regime.

Note that the maximum of the caustic diffraction pattern, essentially given by the Airy function, occurs *before* the caustic point at $\tau = 0$. Soon after the caustic point, we have seen that an Airy signal behaves as $\sim e^{-\tau^{3/2}}$ (cf. Fig. 6), but this over-exponential damping should be replaced by an appropriate ringdown model. This suggests a natural prediction of the model, namely an Airy-relation between the maximum of the waveform and the transition to linearity signaled by the ‘‘merger’’ of the ‘‘fold-caustic’’ stationary points at $\tau = 0$. We are led to the following conjecture for the transition from the merger to the ringdown regime. Several recent investigations have proposed that the ringdown description of the post-

merger signal in terms of damped sinusoidal waveforms can be extended backwards in time towards the merger, even up to the peak of the signal [50, 55] (see also [53, 54]). If the Airy function is indeed an accurate representation of the signal, it suggests then the possibility that the rapid decay of the Airy function after the peak at $\tau \approx -1.019$ might mimic an exponentially damped signal without oscillations. The Airy function decreases monotonically from ≈ 0.5357 at its last peak to ≈ 0.3550 at the caustic $\tau = 0$. This is consistent with the fact that the oscillation frequencies found in [50, 55] are poorly constrained while the damping time is found more accurately and is the dominant effect. On the other hand, the true ringdown phase describing the final remnant black hole is expected to become active only after the transition time $\tau = 0$, which occurs after the peak. The conjecture is that this transition to the ringdown occurs exactly at 1.019 (dimensionless) units of Airy time τ after the peak, namely at the inflection point of the signal when parametrised in the τ time.

V. CONCLUSIONS

Despite the non-linearities of general relativity, binary black hole mergers waveforms are seen to be simple and to have universal properties. These include: i) the increase in amplitude and rapid extinction shortly after the merger, ii) transition from an oscillatory to a damped regime as we cross the merger, and iii) the signal in the merger is seen to be relatively independent of the initial conditions. There exist separate approximation schemes in the inspiral and ringdown regimes which are respectively the post-Newtonian formalism and the framework of black hole perturbation theory. Both these formalisms break down as we go towards the merger. In addition, there are further indications of surprising linearity in the merger phase coming from numerical and observational studies of the ringdown phase.

In this paper the simplicity and universality of the BBH merger waveform are addressed by adopting an approach that simplifies the details of the problem and focuses on its structurally stable aspects, in the spirit of an ‘asymptotic reasoning’ [39, 93]. In particular, we explore the notion that the BBH merger waveform universal features are similar to those found in numerous other natural phenomena, in particular in optical phenomena involving caustics. The framework of singularity theory provides a classification of caustics under the assumption of structural stability, and it also leads to useful expressions of the radiation field in terms of universal functions given by Fresnel integrals for diffraction over the caustics. Specifically, caustics in ‘geometric optics’ represent the “geometrical skeleton supporting the wave flesh” [72], the latter being provided by diffraction in ‘catastrophe optics’. We have discussed the axisymmetric lens and the rainbow as simple examples, and we have seen that the Airy function associated with the fold catastrophe provides an appropriate model for the BBH merger waveform. In particular, the associated Airy equation provides a concrete realization of an effective linearity in this physical transient process.

We have argued that binary black hole merger waveforms

share many of the features of a fold caustic, and we are led to postulate that the merger signal can be described as a reparameterized Airy function $\text{Ai}(\tau(t))$ and a slowly varying amplitude $a(t)$ which remains finite at the merger. At a practical level, our model is a substitution of the conventional reparameterized sine/cosine functions by the reparameterized Airy function. We have proposed a concrete model for the binary merger waveforms given in Eqs. (53) and (54) meant to capture the inspiral-merger transition. This is not (yet) a full fledged inspiral-merger-ringdown model optimized for use in gravitational wave data analysis.

The Airy function has a natural time $\tau = 0$ which can be identified as the merger time which separates the oscillatory and damped regimes. However, the peak of the Airy function occurs *before* $\tau = 0$. Thus, we conjecture that the common identification of the waveform peak with the merger time is incorrect. From this perspective, it is natural to attach an appropriate ringdown model only for $\tau > 0$. This has important consequences, for example, for issues related with black hole spectroscopy and the identification of ringdown overtones. Just like the phase function in the inspiral regime, the reparameterization function $\tau(t)$ and the amplitude $a(t)$ depend on the binary parameters. A first comparison of $\tau(t)$ and $a(t)$ with the leading order post-Newtonian result and with numerically generated merger waveforms can be carried out. However, much more remains to be done; the detailed dependence of $(\tau(t), a(t))$ on the mass ratio, spins, eccentricity will be explored elsewhere, and this will be critical for applications to gravitational wave astronomy and data analysis.

It is important to keep in mind that the ringdown itself is a distinct phenomena and needs to be added separately; it depends on the finite-size nature of the remnant object and is not automatically part of the present catastrophe theory framework built on the basis of rays in geometric optics. Thus, as a first approximation, neutron star mergers share several of these features for the inspiral and merger regimes. However, unlike binary black hole mergers, the remnant is very different. The phenomena which model this regime are clearly very different from what we see in a binary black hole merger remnant. Nevertheless, the Airy function model could be used to link an appropriate post-merger model with the inspiral.

We have argued that similar behavior should be viable in alternate theories of gravity as well. Catastrophe theory contains the notion of structural stability into its fundamental assumptions which means that small perturbations should not lead to qualitatively different behavior. In fact, it can be argued that none of our models of a physical system can be *exactly* true and all of our models should satisfy the notion of structural stability in order to be viable. This leads us to conjecture that binary black hole merger waveforms in alternate theories of gravity should share the features of simplicity and universality seen in standard general relativity. In fact, we can include deviations from standard general relativity: the parameterized post-Newtonian framework (including diffeomorphism invariant metric theories of gravity) results in qualitatively similar behavior for the pre-merger regime. Similarly, it is reasonable to postulate that the remnant black hole in many alternate theories of gravity should approach equilibrium by

emitting damped sinusoidal radiation. We are therefore again led to the transition from an oscillatory to damped behavior, and catastrophe theory again leads to the Airy function model.

While the immediate goals are to develop and test this model with a view towards applications in gravitational wave astronomy, it is plausible that this work will enable us to discover new physical phenomena connected with the merger. One of these is the Arnold exponent β which appears in the scaling laws of the amplitude at the merger. From the Airy function model, we predict that $\beta = 1/6$. An independent confirmation via numerics and/or analytic work (and eventually by observations) would be of interest. More broadly, the ultimate goal is to derive this model from first principles starting with the Einstein equations and following systematic approximation schemes. In this setting it is tantalizing to consider a specially important non-linear generalization of the Airy function, namely the Painlevé-II Transcendent, intimately associated to integrable systems [99]. In particular, it will be of interest to investigate the role of integrability in binary black hole dynamics and its implications for universality [65].

ACKNOWLEDGMENTS

We thank Abhay Ashtekar, Ivan Booth and Andrey Shoom for discussions at the early stage of this project, at the “Focus

Session: Dynamical Horizons, Binary Coalescences, Simulations and Waveform” at the Pennsylvania State University in July 2018. We are especially indebted to Oscar Meneses-Rojas and Ricardo Uribe-Vargas for the patient and generous sharing of their insights in singularity theory. We also thank Bruce Allen, Carlos Barceló, Beatrice Bonga, Sam Dolan, Vincent Lam, Klaas Landsman, Peter Miller, Alex Nielsen, Ariadna Ribes-Metideri, Mikhail Semenov Tian-Shansky, Dhruv Sharma, Carlos F. Sopuerta and Théo Torres for valuable discussions. We also thank the Banff International Research Station for enabling numerous discussions during the Workshop “At the Interface of Mathematical Relativity and Astrophysics”. We are grateful to all participants in this workshop. This work was supported by the French “Investissements d’Avenir” program through project ISITE-BFC (ANR-15-IDEX-03), the ANR “Quantum Fields interacting with Geometry” (QFG) project (ANR-20-CE40-0018-02), the EIPHI Graduate School (ANR-17- EURE-0002) and the Spanish FIS2017-86497-C2-1 project (with FEDER contribution).

-
- [1] F. Pretorius, Evolution of Binary Black Hole Spacetimes, *Phys. Rev. Lett.* **95**, 121101 (2005), arXiv:gr-qc/0507014.
 - [2] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Accurate evolutions of orbiting black-hole binaries without excision, *Phys. Rev. Lett.* **96**, 111101 (2006), arXiv:gr-qc/0511048 [gr-qc].
 - [3] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, Gravitational wave extraction from an inspiraling configuration of merging black holes, *Phys. Rev. Lett.* **96**, 111102 (2006), arXiv:gr-qc/0511103 [gr-qc].
 - [4] M. Boyle *et al.*, The SXS Collaboration catalog of binary black hole simulations, *Class. Quant. Grav.* **36**, 195006 (2019), arXiv:1904.04831 [gr-qc].
 - [5] J. Healy and C. O. Lousto, Third RIT binary black hole simulations catalog, *Phys. Rev. D* **102**, 104018 (2020), arXiv:2007.07910 [gr-qc].
 - [6] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, Georgia Tech Catalog of Gravitational Waveforms, *Class. Quant. Grav.* **33**, 204001 (2016), arXiv:1605.03204 [gr-qc].
 - [7] B. P. Abbott *et al.* (Virgo, LIGO Scientific), Directly comparing GW150914 with numerical solutions of Einstein’s equations for binary black hole coalescence, *Phys. Rev. D* **94**, 064035 (2016), arXiv:1606.01262 [gr-qc].
 - [8] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA), GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run, (2021), arXiv:2111.03606 [gr-qc].
 - [9] A. H. Nitz, S. Kumar, Y.-F. Wang, S. Kastha, S. Wu, M. Schäfer, R. Dhurkunde, and C. D. Capano, 4-OGC: Catalog of gravitational waves from compact-binary mergers, (2021), arXiv:2112.06878 [astro-ph.HE].
 - [10] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA), First joint observation by the underground gravitational-wave detector, KAGRA, with GEO600, (2022), arXiv:2203.01270 [gr-qc].
 - [11] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Directly comparing GW150914 with numerical solutions of Einstein’s equations for binary black hole coalescence, *Phys. Rev. D* **94**, 064035 (2016), arXiv:1606.01262 [gr-qc].
 - [12] L. Blanchet, Gravitational radiation from post-newtonian sources and inspiralling compact binaries, *Liv. Rev. Relat.* **9**, 4 (2006), uRL (cited on 28 January 2008): <http://www.livingreviews.org/lrr-2006-4>.
 - [13] R. A. Porto, The effective field theorist’s approach to gravitational dynamics, *Phys. Rept.* **633**, 1 (2016), arXiv:1601.04914 [hep-th].
 - [14] K. D. Kokkotas and B. G. Schmidt, Quasinormal modes of stars and black holes, *Living Rev. Rel.* **2**, 2 (1999), arXiv:gr-qc/9909058 [gr-qc].
 - [15] E. E. Flanagan and T. Hinderer, Constraining neutron star tidal Love numbers with gravitational wave detectors, *Phys. Rev. D* **77**, 021502 (2008), arXiv:0709.1915 [astro-ph].
 - [16] T. Damour and A. Nagar, Relativistic tidal properties of neutron stars, *Phys. Rev. D* **80**, 084035 (2009), arXiv:0906.0096 [gr-qc].
 - [17] D. Christodoulou, Nonlinear nature of gravitation and gravitational wave experiments, *Phys. Rev. Lett.* **67**, 1486 (1991).
 - [18] J. Frauendiener, Note on the memory effect, *Classical and Quantum Gravity* **9**, 1639 (1992).

- [19] K. S. Thorne, Gravitational-wave bursts with memory: The Christodoulou effect, *Phys. Rev. D* **45**, 520 (1992).
- [20] L. Blanchet and T. Damour, Hereditary effects in gravitational radiation, *Phys. Rev. D* **46**, 4304 (1992).
- [21] M. Hübner, C. Talbot, P. D. Lasky, and E. Thrane, Measuring gravitational-wave memory in the first LIGO/Virgo gravitational-wave transient catalog, *Phys. Rev. D* **101**, 023011 (2020), arXiv:1911.12496 [astro-ph.HE].
- [22] P. D. Lasky, E. Thrane, Y. Levin, J. Blackman, and Y. Chen, Detecting gravitational-wave memory with LIGO: implications of GW150914, *Phys. Rev. Lett.* **117**, 061102 (2016), arXiv:1605.01415 [astro-ph.HE].
- [23] M. Favata, Nonlinear gravitational-wave memory from binary black hole mergers, *Astrophys. J. Lett.* **696**, L159 (2009), arXiv:0902.3660 [astro-ph.SR].
- [24] N. Khera, B. Krishnan, A. Ashtekar, and T. De Lorenzo, Inferring the gravitational wave memory for binary coalescence events, *Phys. Rev. D* **103**, 044012 (2021), arXiv:2009.06351 [gr-qc].
- [25] B. F. Schutz, The Art and science of black hole mergers, in *Conference on Growing Black Holes: Accretion in a Cosmological Context* (2004) arXiv:gr-qc/0410121.
- [26] A. Buonanno and T. Damour, Effective one-body approach to general relativistic two-body dynamics, *Phys. Rev. D* **59**, 084006 (1999), arXiv:gr-qc/9811091.
- [27] A. Nagar *et al.*, Time-domain effective-one-body gravitational waveforms for coalescing compact binaries with nonprecessing spins, tides and self-spin effects, *Phys. Rev. D* **98**, 104052 (2018), arXiv:1806.01772 [gr-qc].
- [28] T. Damour and A. Nagar, The Effective-One-Body Approach to the General Relativistic Two Body Problem (2016) pp. 273–312.
- [29] A. Bohé *et al.*, Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors, *Phys. Rev. D* **95**, 044028 (2017), arXiv:1611.03703 [gr-qc].
- [30] A. Taracchini *et al.*, Effective-one-body model for black-hole binaries with generic mass ratios and spins, *Phys. Rev. D* **89**, 061502 (2014), arXiv:1311.2544 [gr-qc].
- [31] S. Khan, F. Ohme, K. Chatziioannou, and M. Hannam, Including higher order multipoles in gravitational-wave models for precessing binary black holes, *Phys. Rev. D* **101**, 024056 (2020), arXiv:1911.06050 [gr-qc].
- [32] S. Khan, K. Chatziioannou, M. Hannam, and F. Ohme, Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects, *Phys. Rev. D* **100**, 024059 (2019), arXiv:1809.10113 [gr-qc].
- [33] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, First higher-multipole model of gravitational waves from spinning and coalescing black-hole binaries, *Phys. Rev. Lett.* **120**, 161102 (2018), arXiv:1708.00404 [gr-qc].
- [34] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era, *Phys. Rev. D* **93**, 044007 (2016), arXiv:1508.07253 [gr-qc].
- [35] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal, *Phys. Rev. D* **93**, 044006 (2016), arXiv:1508.07250 [gr-qc].
- [36] L. Santamaría *et al.*, Matching post-Newtonian and numerical relativity waveforms: systematic errors and a new phenomenological model for non-precessing black hole binaries, *Phys. Rev. D* **82**, 064016 (2010), arXiv:1005.3306 [gr-qc].
- [37] P. Ajith *et al.*, Inspiral-merger-ringdown waveforms for black-hole binaries with non-precessing spins, *Phys. Rev. Lett.* **106**, 241101 (2011), arXiv:0909.2867 [gr-qc].
- [38] P. Ajith *et al.*, A Template bank for gravitational waveforms from coalescing binary black holes. I. Non-spinning binaries, *Phys. Rev. D* **77**, 104017 (2008), [Erratum: *Phys.Rev.D* **79**, 129901 (2009)], arXiv:0710.2335 [gr-qc].
- [39] R. W. Batterman, *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence* (Oxford University Press, 2001).
- [40] The notion of ‘asymptotic reasoning’ is discussed by Batterman in [39]: “this type of reasoning involves, at its heart, a type of abstraction—a means for ignoring or throwing away various details. [...] I call this kind of reasoning ‘asymptotic reasoning’ “. Slightly rephrased also in Batterman [39]: “scientific understanding often requires methods which eliminate detail and, in some sense, precision [...]. I call these methods ‘asymptotic methods’ and the type(s) of reasoning they involve ‘asymptotic reasoning’.
- [41] A. Buonanno and T. Damour, Transition from inspiral to plunge in binary black hole coalescences, *Phys. Rev. D* **62**, 064015 (2000), arXiv:gr-qc/0001013.
- [42] T. Damour, The problem of motion in Newtonian and Einsteinian gravity., in *Three Hundred Years of Gravitation* (1987) pp. 128–198.
- [43] R. H. Price and J. Pullin, Colliding black holes: The Close limit, *Phys. Rev. Lett.* **72**, 3297 (1994), arXiv:gr-qc/9402039.
- [44] R. J. Gleiser, C. O. Nicasio, R. H. Price, and J. Pullin, Colliding black holes: How far can the close approximation go?, *Phys. Rev. Lett.* **77**, 4483 (1996), arXiv:gr-qc/9609022.
- [45] G. Khanna, J. G. Baker, R. J. Gleiser, P. Laguna, C. O. Nicasio, H.-P. Nollert, R. Price, and J. Pullin, Inspiralling black holes: The Close limit, *Phys. Rev. Lett.* **83**, 3581 (1999), arXiv:gr-qc/9905081.
- [46] C. F. Sopuerta, N. Yunes, and P. Laguna, Gravitational Recoil from Binary Black Hole Mergers: The Close-Limit Approximation, *Phys. Rev. D* **74**, 124010 (2006), [Erratum: *Phys.Rev.D* **75**, 069903 (2007), Erratum: *Phys.Rev.D* **78**, 049901 (2008)], arXiv:astro-ph/0608600.
- [47] C. F. Sopuerta, N. Yunes, and P. Laguna, Gravitational recoil velocities from eccentric binary black hole mergers, *Astrophys. J. Lett.* **656**, L9 (2007), arXiv:astro-ph/0611110.
- [48] J. G. Baker, B. Bruegmann, M. Campanelli, C. O. Lousto, and R. Takahashi, Plunge wave forms from inspiralling binary black holes, *Phys. Rev. Lett.* **87**, 121103 (2001), arXiv:gr-qc/0102037.
- [49] J. G. Baker, M. Campanelli, and C. O. Lousto, The Lazarus project: A Pragmatic approach to binary black hole evolutions, *Phys. Rev. D* **65**, 044001 (2002), arXiv:gr-qc/0104063.
- [50] M. Giesler, M. Isi, M. A. Scheel, and S. Teukolsky, Black Hole Ringdown: The Importance of Overtones, *Phys. Rev.* **X9**, 041060 (2019), arXiv:1903.08284 [gr-qc].
- [51] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, Testing the no-hair theorem with GW150914, *Phys. Rev. Lett.* **123**, 111102 (2019), arXiv:1905.00869 [gr-qc].
- [52] M. Okounkova, Revisiting non-linearity in binary black hole mergers, (2020), arXiv:2004.00671 [gr-qc].
- [53] C. D. Capano, M. Cabero, J. Westerweck, J. Abedi, S. Kasta, A. H. Nitz, A. B. Nielsen, and B. Krishnan, Observation of a multimode quasi-normal spectrum from a perturbed black hole, (2021), arXiv:2105.05238 [gr-qc].

- [54] R. Cotesta, G. Carullo, E. Berti, and V. Cardoso, On the detection of ringdown overtones in GW150914, (2022), arXiv:2201.00822 [gr-qc].
- [55] M. Isi and W. M. Farr, Revisiting the ringdown of GW150914, (2022), arXiv:2202.02941 [gr-qc].
- [56] J. L. Jaramillo, R. P. Macedo, P. Moesta, and L. Rezzolla, Black-hole horizons as probes of black-hole dynamics I: post-merger recoil in head-on collisions, *Phys.Rev.* **D85**, 084030 (2012), arXiv:1108.0060 [gr-qc].
- [57] J. L. Jaramillo, R. P. Macedo, P. Moesta, and L. Rezzolla, Black-hole horizons as probes of black-hole dynamics II: geometrical insights, *Phys.Rev.* **D85**, 084031 (2012), arXiv:1108.0061 [gr-qc].
- [58] J. Jaramillo, R. Macedo, P. Moesta, and L. Rezzolla, Towards a cross-correlation approach to strong-field dynamics in Black Hole spacetimes, *AIP Conf.Proc.* **1458**, 158 (2011), arXiv:1205.3902 [gr-qc].
- [59] A. Gupta, B. Krishnan, A. Nielsen, and E. Schnetter, Dynamics of marginally trapped surfaces in a binary black hole merger: Growth and approach to equilibrium, *Phys. Rev.* **D97**, 084028 (2018), arXiv:1801.07048 [gr-qc].
- [60] V. Prasad, A. Gupta, S. Bose, B. Krishnan, and E. Schnetter, News from horizons in binary black hole mergers, (2020), arXiv:2003.06215 [gr-qc].
- [61] D. A. B. Iozzo *et al.*, Comparing Remnant Properties from Horizon Data and Asymptotic Data in Numerical Relativity, *Phys. Rev. D* **103**, 124029 (2021), arXiv:2104.07052 [gr-qc].
- [62] A. Ashtekar, N. Khera, M. Kolanowski, and J. Lewandowski, Non-expanding horizons: multipoles and the symmetry group, *JHEP* **01**, 028, arXiv:2111.07873 [gr-qc].
- [63] A. Ashtekar, N. Khera, M. Kolanowski, and J. Lewandowski, Charges and fluxes on (perturbed) non-expanding horizons, *JHEP* **02**, 066, arXiv:2112.05608 [gr-qc].
- [64] L. Sberna, P. Bosch, W. E. East, S. R. Green, and L. Lehner, Nonlinear effects in the black hole ringdown: absorption-induced mode excitation (2021), arXiv:2112.11168 [gr-qc].
- [65] J. L. Jaramillo and B. Krishnan, Painlevé-II approach to binary black hole merger dynamics: universality from integrability, In preparation.
- [66] S. Borhanian, K. G. Arun, H. P. Pfeiffer, and B. S. Sathyaprakash, Comparison of post-Newtonian mode amplitudes with numerical relativity simulations of binary black holes, *Class. Quant. Grav.* **37**, 065006 (2020), arXiv:1901.08516 [gr-qc].
- [67] Caustic phenomena are, of course, not exclusive from optics and occur generically in the high-frequency (geometric optics) limit of wave theories, such as acoustics, seismic fields... For concreteness, we focus on the comparison with optics.
- [68] K. Thorne and R. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics* (Princeton University Press, 2017).
- [69] J. Adam, *Rays, Waves, and Scattering: Topics in Classical Mathematical Physics*, Princeton Series in Applied Mathematics (Princeton University Press, 2017).
- [70] M. Berry, Waves and thom's theorem, *Advances in Physics* **25**, 1 (1976), <https://doi.org/10.1080/00018737600101342>.
- [71] M. Berry and C. Upstill, *Iv catastrophe optics: Morphologies of caustics and their diffraction patterns* (Elsevier, 1980) pp. 257 – 346.
- [72] Y. Kravtsov and Y. Orlov, *Caustics, Catastrophes and Wave Fields*, Springer Series on Wave Phenomena (Springer Berlin Heidelberg, 2012).
- [73] R. Gilmore, *Catastrophe Theory for Scientists and Engineers*, Dover books on advanced mathematics (Dover Publications, 1993).
- [74] T. Poston and I. Stewart, *Catastrophe Theory and Its Applications*, Dover books on mathematics (Dover Publications, 1996).
- [75] In geometric optics one usually starts actually from the action function, whose gradient provides the rays' wave vectors. This is in general a multivalued function on D , the different branches corresponding to the folding of M in $S \times D$. However in catastrophe theory this 'multi-valuedness' is by-passed by introducing the generating function Φ depending on the additional state variables to render it a single-valued function, in particular non-singular in its arguments.
- [76] This intermediate layer provides another instance of 'asymptotic reasoning' in the sense of Batterman [39, 93], where full details are sacrificed in order to unveil the underlying relevant mechanism.
- [77] J. J. Duistermaat, Oscillatory integrals, lagrange immersions and unfolding of singularities, *Communications on Pure and Applied Mathematics* **27**, 207 (1974), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160270205>.
- [78] V. Guillemin and S. Sternberg, *Symplectic Techniques in Physics* (Cambridge University Press, 1990).
- [79] C. Chester, B. Friedman, and F. Ursell, An extension of the method of steepest descents, *Mathematical Proceedings of the Cambridge Philosophical Society* **53**, 599–611 (1957).
- [80] V. P. Maslov, V. I. Arnol'd, and V. C. Bouslaev, *Théorie des perturbations et méthodes asymptotiques*, Etudes mathématiques (Dunod, Paris, 1972) trans. from the Russian.
- [81] I. KRAVTSOV, Two new asymptotic methods in the theory of wave propagation in inhomogeneous media (asymptotic methods in wave propagation theory in inhomogeneous media, reviewing solutions in acoustics and radio physics), *Soviet Physics-Acoustics* **14**, 1 (1968).
- [82] G. Wassermann, *Stability of Unfoldings*, Lecture Notes in Mathematics (Springer, 1974).
- [83] T. Bröcker and L. Lander, *Differentiable Germs and Catastrophes*, Differentiable Germs and Catastrophes (Cambridge University Press, 1978).
- [84] T. Torres, M. Lloyd, S. R. Dolan, and S. Weinfurter, Wave focusing by submerged islands and gravitational analogues, arXiv preprint arXiv:2202.05926 (2022).
- [85] We would like to highlight the remarkable timing between the notable work [84] and the present one, where uniform asymptotic approximations for caustic diffraction have been implemented in a gravitational physics setting, in both cases largely building on pioneering Berry's work [70, 71, 100].
- [86] A better approximation would be obtained by expanding the third term in (26) around the critical points, keeping higher-order terms. This leads to an asymptotic series (see [79] for the fold-catastrophe).
- [87] Starting from Eq. (22), it follows the relation
- $$\nabla_{R_\alpha} \Phi dR_\alpha = \nabla_{r_\alpha} \phi dr_\alpha . \quad (58)$$
- Then we can estimate $g(\mathbf{r}, \mathbf{X})$ in Eq. (25) with the approximation (27) and, by expanding (58) to second order, that provides
- $$\left| \frac{d\mathbf{R}}{dr} \right| (\mathbf{X}) = \left(\frac{\text{Hess}(\phi)(\mathbf{r}(\mathbf{x}(\mathbf{X})))}{\text{Hess}(\Phi)(\mathbf{R}(\mathbf{X}))} \right)^{\frac{1}{2}}, \quad (59)$$
- and the following $n + 1$ equations at the critical points
- $$\left(\frac{\text{Hess}(\phi)(\mathbf{r}_i(\mathbf{x}(\mathbf{X})))}{\text{Hess}(\Phi)(\mathbf{R}_i(\mathbf{X}))} \right)^{\frac{1}{2}} a(\mathbf{R}_i(\mathbf{X}), \mathbf{X}) \quad (60)$$
- $$= g_0(\mathbf{x}(\mathbf{X})) + g_1(\mathbf{x}(\mathbf{X})) \cdot \nabla_{\mathbf{x}} \phi(\mathbf{r}_i(\mathbf{x}(\mathbf{X})), \mathbf{x}(\mathbf{X})), \quad (61)$$
- fix $g_0(\mathbf{x}(\mathbf{X}))$ and the n -vector $g_1(\mathbf{x}(\mathbf{X}))$, for each \mathbf{X} .

- [88] K. W. Ford and J. A. Wheeler, Semiclassical description of scattering, *Annals of Physics* **7**, 259 (1959).
- [89] O. Meneses-Rojas and et al., In preparation (2022).
- [90] M. Berry, Focusing and twinkling: critical exponents from catastrophes in non-gaussian random short waves, *Journal of Physics A: Mathematical and General* **10**, 2061 (1977).
- [91] Such self-similar behavior captured by Eq. (30) close to the caustic could play in the interplay between self-similarity and phase transitions discussed in [65] for BBH merger waveforms from a PDE perspective.
- [92] In a further step, it will be of interest to explore the relation of R and r with physical system parameters, accounting for the BBH caustic description from first principles in terms of the evolution of the two orbiting black holes in an appropriate point-particle modeling of the BBH dynamics [89].
- [93] R. W. Batterman, ‘into a mist’: Asymptotic theories on a caustic, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **28**, 395 (1997).
- [94] A. Erdélyi, *Asymptotic expansions*, 3 (Courier Corporation, 1956).
- [95] M. Taylor, Airy functions and airy quotients, <https://mtaylor.web.unc.edu/wp-content/uploads/sites/16915/2018/04/airyf.pdf>.
- [96] It should be kept in mind that this corresponds to the transitional approximation. We still have the flexibility to include the derivative of the Airy function (actually another appropriate diffraction integral $J(\mathbf{x})$ in Eq. (29)) and a time dependent phase as given in Eq. (37), at the price of including additional control parameters, e.g. the full spacetime coordinates x^a .
- [97] It would also be consistent to refer to $\tau = 0$ as the “coalescence time”, since it corresponds to $t = t_c$ in Post-Newtonian expansions. We refer however to it as the “merger time” in the present caustic model, since it is the only time naturally related to the merger in the present diffraction caustic framework.
- [98] This raises the possibility of exploring the possibility of τ and/or a as appropriate “state variables” in the BBH fold caustic description, in the spirit of Fig. 5.
- [99] R. Conte and M. Musette, *The painlevé handbook*, (2008).
- [100] M. Berry, *A Half-Century of Physical Asymptotics and Other Diversions: Selected Works by Michael Berry* (World Scientific Publishing Company, 2017).

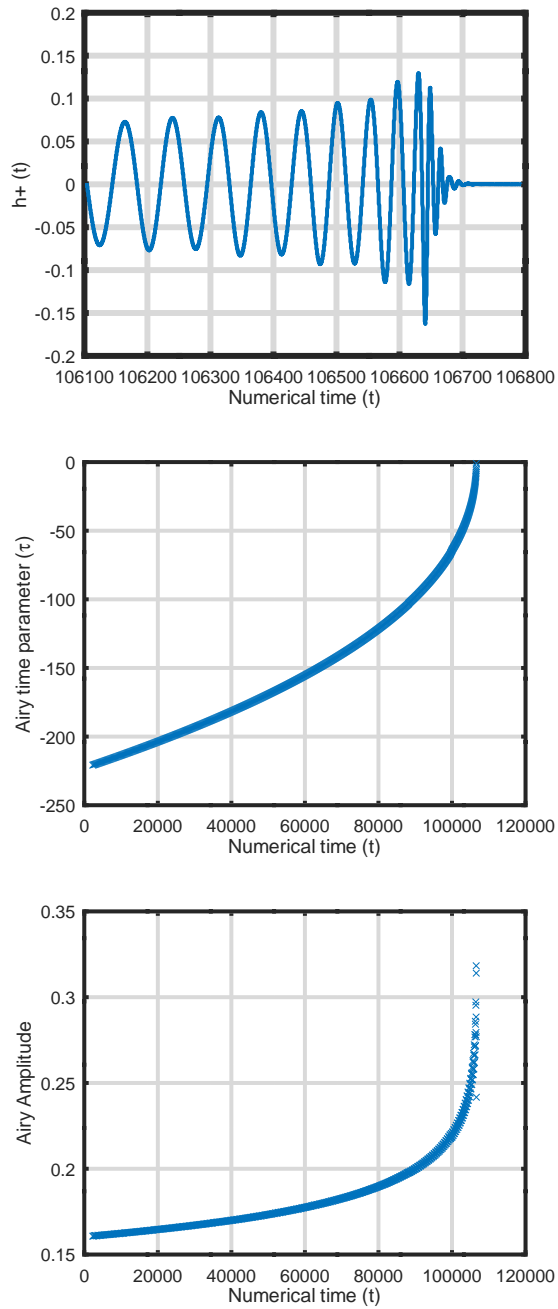


FIG. 7. The Airy reparametrization function and the Airy amplitude for a particular NR waveform from the SXS catalog. The first panel shows the late part of the waveform. The second panel is the time mapping function $\tau(t)$, and the last panel is the Airy amplitude $a(t)$ appearing in Eqs. (53) and (53) following the procedure as described in the text. Despite appearances, the values of τ in the second panel terminate at the Airy peak time $\tau = -1.019$, and do not reach $\tau = 0$.