Wave Propagation in Gravitational Systems: Completeness of Quasinormal Modes

E.S.C. Ching[1], P.T. Leung[1], W.M. Suen[2,3] and K. Young[1]

[1] *Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong*

[2] *McDonnell Center for the Space Sciences, Department of Physics, Washington University, St Louis, MO 63130*

[3] *National Center for Supercomputing Applications, University of Illinois, Urbana, IL, 61801*

(July 17, 1995)

Abstract

The dynamics of relativistic stars and black holes are often studied in terms of the quasinormal modes (QNM's) of the Klein-Gordon (KG) equation with different effective potentials $V(x)$. In this paper we present a systematic study of the relation between the structure of the QNM's of the KG equation and the form of $V(x)$. In particular, we determine the requirements on $V(x)$ in order for the QNM's to form complete sets, and discuss in what sense they form complete sets. Among other implications, this study opens up the possibility of using QNM expansions to analyse the behavior of waves in relativistic systems, even for systems whose QNM's do *not* form a complete set. For such systems, we show that a complete set of QNM's can often be obtained by introducing an infinitesimal change in the effective potential.

PACS numbers: 04.30.-w, 95.30.Sf, 04.30.Nk, 04.30.Db

Typeset using REVTEX

I. INTRODUCTION

Gravitational waves have been of theoretical interest since the appearance of the general theory of relativity, and of experimental interest for several decades. The generation and propagation of gravitational waves are often discussed in terms of a source region, a strong-field and a weak-field near zone, and a local and distant wave zone [1]. For the purpose of this paper, a simpler point of view will be useful for understanding: (i) a source region where the radiation is generated, (ii) an intervening region with a non-trivial geometry of the background spacetime, in which the propagation of waves is significantly different from that in flat space, and (iii) an "outside" region where the spacetime can be taken to be flat, in which the observer is located. A close analogy is a molecule or molecules [region (i)] radiating in an optical cavity [region (ii)]. The cavity significantly modifies the character of the electromagnetic radiation observed outside [region (iii)].

The interest of this paper is on the effects of the intervening nontrivial spacetime, i.e., the "cavity", in gravitational systems. For example, consider radiation from a compact astrophysical configuration, e.g., a binary neutron star system, embedded in a galaxy. The source region is the binary system, the "cavity" is the galaxy, and the rest the "outside". The presence of a "cavity" in principle could have a significant influence on the radiation observed from the outside. The issue of a recent debate [2] is whether the "cavity", in this case, the galaxy, can suppress, by many orders of magnitude, the emergence of radiation generated by the source [3]. Another example is a particle radially plunging into a black hole of mass M. In this case the "source region" and the "cavity" have the same size (as is true for a laser cavity). The "source" is significant only within $\sim 10M$ of the plunge, and the effective potential (in the perturbation theory of a Schwarzschild hole [4]) representing the "cavity" is also significant only within $\sim 10M$. It is well-known that in this case the "cavity" modifies the observed radiation significantly.

Radiation from optical cavities is often discussed in terms of the "modes" of the cavity. Because the waves escape to infinity, the total energy in the finite part of the system decreases with time. Thus, these "modes" are characterized by complex frequencies ω_j, with $\mathrm{Im}\,\omega_j < 0$. The corresponding eigenfunctions are defined by the outgoing wave condition at infinity. These "modes" are in fact quasinormal modes (QNM's). The observation of the QNM's from the outside gives directly information on the spatial structure of the cavity (but not the structure of the source, unless a particular model is assumed); e.g., in obvious notation, $\omega_j \sim j\pi c/L$, where L is the length of a simple 1-dimensional optical cavity.

The same discussion extends to the gravitational case. The QNM's of black holes [4] and relativistic stars [5] have been subjects of much study. In numerical simulations, it is often found [6] that the radiation observed in many black hole processes is dominated by the QNM's of the background spacetime associated with the hole. In view of the optical analogy, this is not surprising — the distant observer sees only the QNM's of the laser cavity, but *not* the details of the source. The fact that gravitational QNM's tend to be more "leaky" does not change this qualitative understanding.

It is therefore particularly exciting that gravitational QNM radiation may be observed by LIGO, VIRGO, and other detectors [7] in the next decade. These observations will provide the possibility of "seeing" directly interesting spacetime structures of various gravitational systems, e.g., the spacetime around a black hole — in much the same way as the spectrum of a laser permits a distant observer to infer the geometry of the cavity.

Although the QNM's of black holes are quite well studied, the general behavior of QNM's in more complicated gravitational systems has not received much attention. For example, in view of the possible detection of gravitational waves, one would like to know: How much would the QNM frequencies of a black hole be shifted by the astrophysical environment that it is in, e.g., an accretion disk or the host galaxy? Can one calculate the shift perturbatively? How much would each QNM be excited in an astrophysical process, e.g., a star plunging into a massive hole? There are many such questions of both theoretical and observational interest.

In an ongoing project, we study the general properties of QNM's in gravitational systems. In a recent letter [8], we outlined the results obtained in studying the QNM's of the Klein-

Gordon (KG) equation

$$D\phi \equiv \left[\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} + V(x)\right]\phi(x,t) = 0, \tag{1.1}$$

which is often used to describe wave propagation in curved space, with $V(x)$ describing the scattering of waves by the geometry. In this paper, we provide the details of the results reported in [8]. We also present other related results and discuss the implications.

For a set of modes, the first question is whether they are complete. Completeness can mean several different things, and we shall discuss the relations among the various senses of completeness in Sec. III; for the moment we take the simple point of view that the QNM's are complete if the evolution of the wavefunction can be expressed, in some domain of time and space, as a sum

$$\sum_j a_j f_j(x) e^{-i\omega_j t}, \tag{1.2}$$

where $f_j(x)$ are eigenfunctions with complex frequencies ω_j. In particular, if this holds at $t = 0$, then an arbitrary function of x within a certain domain should be expressible as

$$\sum_j a_j f_j(x). \tag{1.3}$$

Both of these properties can be discussed in terms of a Green's function $G(x, y; t)$ for propagating a source at y to an observation point at x in time t. Our purpose is to elucidate under what conditions would completeness hold. This problem is nontrivial because usual proofs of completeness for the normal modes of conservative systems rely crucially on hermiticity. It is therefore particularly interesting that Price and Husain [9] were able to give a model of relativistic stellar oscillations in which the QNM's do form a complete set. It was suggested [9] that the completeness of the QNM's stems either from the equality of the damping times of all the QNM's, or from the absence of dispersion and backscatter in that model. We shall see in this paper that these properties are not the controlling factors of completeness.

There is a limit in which QNM's should behave trivially, like the normal modes of hermitian systems, namely the limit in which the leakage goes to zero; mathematically this means that the quality factor $Q_j \simeq \mathrm{Re}\ \omega_j/(2\ |\mathrm{Im}\ \omega_j\ |) \to \infty$. In this limit, the QNM's become normal modes and are complete. In the present paper we study the completeness of gravitational QNM's whose quality factors are *not* large, and we are interested in results that survive to all orders in the leakage, which may be characterized by $1/Q_j$. We show that the QNM's are complete in a broad class of models quite independent of leakage, provided the following three conditions hold:

(i) The effective potential $V(x)$ in (1.1) is everywhere finite, and vanishes sufficiently rapidly, in a sense to be defined in Sec. II, as $|x| \to \infty$;

(ii) there are spatial discontinuities in $V(x)$ demarcating the boundaries of a "cavity"; and

(iii) consideration is limited to certain domains of space and time, which will be spelt out below.

It is useful to comment heuristically on these conditions at the outset. If (i) does not hold, i.e., if $V(x)$ has a significant tail at large x, then it is possible for a disturbance originating at y to propagate to a large $x' \gg x$, and be scattered back to the observation point x. The time taken would be $O(x')$, and since x' can be arbitrarily large, this could lead to a late time tail in the wavefunction; crudely speaking this would be $t^{-\alpha}$ if asymptotically $V(x) \sim x^{-\alpha}$ (though this is modified in the presence of a centrifugal barrier [10]). Such a late time tail would not be described by a sum of discrete QNM's, which must behave exponentially at large times. Thus, the QNM's cannot be complete for $V(x)$ having a tail. The late time tail, and its relationship with the spatial asymptotics of $V(x)$, is now thoroughly understood [10].

A discrete sum over QNM's (like a Fourier series rather than a Fourier integral) cannot possibly be complete over all space; completeness, if it holds at all, can only be limited to the inside of a finite interval. The need for a natural demarcation of this interval "explains" why discontinuities are needed. In the model of [9] there is such a discontinuity in the speed of wave propagation at the model stellar surface. This also explains why the completeness relation (1.3) must be limited spatially.

The need to also limit the time domain in some cases can be understood from the following example. Suppose the observation point is far from the source. Then there will be transients propagating directly from the source point y to the observation point x (i.e., the high frequency components), without being much affected by the potential; these have nothing to do with the QNM's. Only past a certain finite prompt time $t_p = O(x)$, when these transients have passed, would the QNM's be complete. This case is most relevant for the discussion of gravitational waves seen by a distant observer, and numerical experiments indeed show that quasinormal ringing dominates only after the passage of transients.

It should be emphasized that the broad class of models satisfying the above conditions and hence exhibiting completeness allow dispersion, backscatter, as well as differences in the damping times of the QNM's, showing that these are not essential ingredients. This broad class of models also allow the leakage to be not small; in other words, our results are valid to all orders in $1/Q_j$.

There is one final technical hedge. The QNM sums are in fact often divergent series, but converge to the correct answer if regularized in a standard way. The simplest statement of regulation is: keep only modes with Re $\omega_j > 0$ by symmetry argument, then replace all times t by $t - i\tau$, with $\tau \rightarrow 0^+$. We shall come to the need for this in Sec. II, and we shall also demonstrate, through a numerical example in Sec. IV, that the regulated series is useful in practice.

The general need for regularization is a major difference between the completeness problem of the KG equation (1.1) and the wave equation

$$\left[n^2(z) \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2} \right] \psi(z,t) = 0. \tag{1.4}$$

where $n(z)$ may be regarded as a position-dependent refractive index in a scalar model of electromagnetism. We have previously studied the completeness of QNM's in the wave equation case [11]. In [9], the stellar oscillation model is described by (1.4), with $n(z) = 1$ for

$0 \leq z \leq L$ representing the interior of the star, and $n(z)$ a non-zero constant for $L < z < \infty$ representing space exterior to the star. The boundary conditions are $\psi(z = 0) = 0$ and outgoing waves at infinity.

In this paper we work in terms of the KG equation (1.1) since it is more realistic as a model for wave propagation in curved space, and has been much studied for this purpose. For example, in a static, spherically symmetric spacetime,

$$ds^2 = g_{tt}(r)dt^2 + g_{rr}(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{1.5}$$

a KG scalar field Φ can be expressed as

$$\Phi = \sum_{lm} \frac{1}{r}\phi_{lm}(x,t)Y_{lm}(\theta,\varphi). \tag{1.6}$$

The evolution of each $\phi_{lm}(x,t)$ is given by (1.1) with the effective potential

$$V(x) = -g_{tt}\frac{l(l+1)}{r^2} - \frac{1}{2r}\frac{g_{tt}}{g_{rr}}(\frac{1}{g_{tt}}\frac{\partial g_{tt}}{\partial r} - \frac{1}{g_{rr}}\frac{\partial g_{rr}}{\partial r}). \tag{1.7}$$

The variable x in the KG equation is related to the circumferential radius r by

$$x = \int \sqrt{-\frac{g_{rr}}{g_{tt}}} dr. \tag{1.8}$$

The particular case of (1.1) arising from a Schwarzschild spacetime is well studied. The Maxwell field and the linearized gravitational waves satisfy the same equation with slightly different potentials $V(x)$ [4].

The wave equation (1.4) is related to the KG equation (1.1) by a transformation:

$$\begin{aligned} dx/dz &= n(z), \\ \psi &= n^{-1/2}\phi, \end{aligned} \tag{1.9}$$

with the potential related to $n(z)$ by

$$V = (2n^3)^{-1}(d^2n/dz^2) - (3/4n^4)(dn/dz)^2 . \tag{1.10}$$

Hence the results obtained in one case can be restated in the other, up to possible complication coming from regularization, which is investigated in this paper. As we shall see later, discontinuity is a crucial feature in the consideration. We note that $V(x)$ with a discontinuity in its p-th derivative maps into $n(z)$ with a discontinuity in its $(p+2)$-th derivative.

Our results for discontinuous potentials, in addition to clarifying the conditions for completeness, are significant in two ways. First, they are directly applicable to physical models with a discontinuity, e.g., a stellar surface as in the model of [9]. Second, any smooth potential can be approximated, to arbitrary precision, by a discontinuous one. In fact, any numerical scheme that employs finite spatial differencing has in effect replaced $V(x)$ by one with a discontinuity in a high order derivative. Thus the complete set of QNM's of the discontinuous system can be used as an effective calculational tool for the original system. Further remarks on this aspect are given in Sec. V, while detailed studies along this line will be reported elsewhere.

In Sec. II we develop the formalism via the Green's function in the complex frequency plane, and demonstrate that for a class of discontinuous potentials V, the Green's function is given by a sum of QNM's. The conceptual analogy with optical cavities extends to a close parallel in the mathematical treatment with the optical case [11], but there are technical differences which are stated in this paper. Sec. III spells out more clearly the different senses of completeness and also discusses briefly the extension to KG equation with a non-zero mass. In Sec. IV, we verify numerically the completeness relation in a model problem. The paper ends with some discussions and conclusion.

II. THE GREEN'S FUNCTION

A. Green's function in frequency plane

We model the propagation of waves in curved space by the KG equation (1.1). The potential $V(x)$ is assumed to be bounded and positive, and represents the scattering of waves by the background. Generically, we take $V_o \equiv \lim_{x\to\infty} V(x) = 0$. The case $V_o \neq 0$, which represents a massive field, will be discussed briefly in Sec. III. The spatial coordinate x often represents a radial variable [cf. (1.8)], so we shall first consider a half-line problem ($x \geq 0$) with the regular boundary condition $\phi(x=0,t) = 0$. The full line problem ($-\infty < x < \infty$), e.g., x being the tortoise coordinate in the Schwarzschild case, will be studied later. The QNM's are eigen-solutions $\phi(x,t) = \tilde{\phi}(x)e^{-i\omega t}$ of (1.1), where $\tilde{\phi}$ satisfies

$$\tilde{D}\tilde{\phi} \equiv [-\omega^2 - \frac{\partial^2}{\partial x^2} + V(x)]\tilde{\phi}(x) = 0, \tag{2.1}$$

with $\tilde{\phi}(x=0) = 0$ and the outgoing wave boundary condition at infinity. The eigenfunctions and eigenvalues are defined as $\tilde{\phi}(x) = f_j(x)$ and $\omega = \omega_j$.

The Green's function $G(x,y;t)$ for the system is defined by $DG = \delta(x-y)\delta(t)$ with the initial condition $G = 0$ for $t \leq 0$ and the boundary conditions that (i) $G = 0$ for either $x = 0$ or $y = 0$, and (ii) the outgoing wave condition as either $x \to \infty$ or $y \to \infty$. The corresponding Green's function in the frequency domain is

$$\tilde{D}\tilde{G}(x,y;\omega) = \delta(x-y). \tag{2.2}$$

The strategy is to express G in terms of $\tilde{G}$, and attempt to close the contour by a large semicircle in the lower half ω plane.

Introduce two auxiliary functions $f(\omega,x)$ and $g(\omega,x)$, which are the solutions to the homogeneous time-independent KG equation $\tilde{D}f(\omega,x) = \tilde{D}g(\omega,x) = 0$ with the boundary conditions $f(\omega,x=0) = 0$; $f'(\omega,x=0) = 1$ [12] and $\lim_{x\to\infty}[g(\omega,x)\exp(-i\omega x)] = 1$. With these auxiliary functions the Green's function is then given by

$$\tilde{G}(x,y;\omega) = f(\omega,x)g(\omega,y)/W(\omega) \quad \text{for} \quad 0<x<y,$$
$$= f(\omega,y)g(\omega,x)/W(\omega) \quad \text{for} \quad 0<y<x, \tag{2.3}$$

where the Wronskian $W(\omega) = g(\omega,x)f'(\omega,x) - f(\omega,x)g'(\omega,x)$ $(' = d/dx)$ is independent of x. $\tilde{G}(x,y;\omega)$, of course, may be singular in ω at the singularities of f and g; otherwise, it is analytic except at zeros of $W(\omega)$. At these zeros f and g are proportional to each other, so each of them satisfies the regular boundary condition at $x=0$, *and* the outgoing wave boundary condition as $x \to \infty$. Hence these frequencies are exactly the QNM frequencies ω_j and $f(\omega_j,x) = f_j(x)$. We shall further assume for simplicity that these zeros are simple [13], so that the residues of $\tilde{G}(x,y;\omega)$ at these zeros are given by

$$K_j = f(\omega_j,x)g(\omega_j,y)/[\partial W(\omega=\omega_j)/\partial\omega]. \tag{2.4}$$

Multiple zeros can be handled readily. To study the physical meaning of the denominator $\partial W/\partial\omega$, start with the defining equation for $f(\omega_j,x)$ and $g(\omega,x)$. The usual manipulations lead to

$$(\omega^2-\omega_j^2)\int_0^X dx\ f(\omega_j,x)g(\omega,x) = g(\omega,x)f'(\omega_j,x) - g'(\omega,x)f(\omega_j,x)|_0^X, \tag{2.5}$$

where the integral is taken along any contour from $x=0$ to $x=X$. Since both $f(\omega_j,x)$ and $g(\omega,x)$ are outgoing waves at $x=X$, the right-hand side of (2.5) becomes

$$i(\omega_j-\omega)f(\omega_j,X)g(\omega,X) - [g(\omega,0)f'(\omega_j,0) - g'(\omega,0)f(\omega_j,0)], \tag{2.6}$$

as $X\to\infty$. Differentiating (2.5) with respect to ω and taking $\omega\to\omega_j$ then gives

$$\lim_{X\to\infty}\int_0^X dx\ f(\omega_j,x)g(\omega_j,x) + \frac{i}{2\omega_j}f(\omega_j,X)g(\omega_j,X) = -\frac{1}{2\omega_j}\frac{\partial W(\omega=\omega_j)}{\partial\omega}. \tag{2.7}$$

Since f and g are proportional to each other at these poles, the residues can be expressed in terms of a generalized norm of the QNM's

$$\ll f_j \mid f_j \gg\ \equiv \lim_{X\to\infty}\int_0^X dx\ f(\omega_j,x)f(\omega_j,x) + \frac{i}{2\omega_j}f(\omega_j,X)f(\omega_j,X) \tag{2.8}$$

as follows

$$K_j = -\frac{f_j(x) f_j(y)}{2\omega_j \ll f_j \mid f_j \gg}. \tag{2.9}$$

The norm (2.8) has been introduced in other contexts [11,14,15] and has the following significant properties:

(i) it invovles f^2 rather than $|f|^2$, and is therefore in general complex;

(ii) there is a surface term;

(iii) although each of the two terms on the right of (2.8) does not have a limit, the limit exists for the sum; and

(iv) it reduces to the usual norm if the leakage goes to zero.

The property (iii) is easily verified by differentiating the right hand side of (2.8) with respect to X, and using the outgoing wave condition, which also applies to f at $\omega = \omega_j$. The appearance of the surface term is closely connected with the fact that the momentum operator is not hermitian. While (2.8), with the integral taken along the real axis, is often convenient for actual evaluation, we may choose a deformed contour L running from $x = 0$ to $x = -\infty$. The first part $L_1 = (0, a)$, and the second part L_2 is shown in Fig. 1. On L_2, $f(\omega_j, x)$ is defined as the outside solution (which is analytic in x for $x > a$), analytically continued to the complex plane. Along this contour, $f(\omega_j, x) \to 0$ as $x \to -\infty$, and we get, more compactly

$$\ll f_j \mid f_j \gg \equiv \int_L dx f(\omega_j, x)^2 dx, \tag{2.10}$$

making the formal analogy to the hermitian case more transparent. However, analytic continuation out into the complex x plane is not strictly necessary if the surface term is retained, as in (2.8).

Next write G in terms of $\tilde{G}$; then upon closure of the contour in the lower half ω plane, one sees that

$$G(x, y; t) = \frac{i}{2} \sum_j \frac{f_j(x) f_j(y) e^{-i\omega_j t}}{\omega_j \ll f_j \mid f_j \gg} + I_c + I_s. \tag{2.11}$$

In (2.11), the sum comes from the zeros of the Wronskian $W(\omega)$ inside the semicircle, I_c is the integral along a semicircle at infinity, and I_s comes from the singularities of $f(\omega, x)$ and $g(\omega, x)$ (see Fig. 2). The crux of the proof of completeness then lies in (i) determining the conditions on $V(x)$ under which f and g have no singularities in the ω plane, in which case $I_s = 0$; and (ii) showing (in the next subsection) that I_c vanishes if there is a discontinuity in the potential $V(x)$ at some $x = a > 0$. Under these conditions, the Green's function can be represented exactly in terms of the QNM's, which then establishes these QNM's as a sufficient basis for discussing the dynamics. The behavior is a discrete sum of exponentials, and in particular, there would be no power-law dependence at large times.

Since QNM's appear in pairs at $\omega = \omega_j$ and $\omega = -\omega_j^*$, it is convenient to rewrite the QNM sum (when $I_s = I_c = 0$) as

$$G(x, y; t) = \text{Re}\left[i \sum_{j>0} \frac{f_j(x) f_j(y) e^{-i\omega_j t}}{\omega \ll f_j | f_j \gg} \right], \tag{2.12}$$

where the notation $j > 0$ is a shorthand for Re $\omega_j > 0$. This latter form is somewhat more convenient for regularization when the sum diverges. The corresponding statements for $\dot{G}(x, y; t)$, and consequently a QNM representation of $\delta(x - y) = \dot{G}(x, y; t = 0^+)$, follow relatively simply and will be discussed in Sec. III.

Now to determine the conditions under which $f(\omega, x)$ and $g(\omega, x)$ have no singularities in ω, we appeal to well known results in the quantum theory of scattering. The defining equation for f and g is identical to the Schrödinger equation for a particle with mass $= 1/2$ and energy $= \omega^2$ moving in a potential well $V(x)$. It has been proved [16] that f and g are analytic functions of ω if the potential is bounded and "has no tail", in the sense that

$$\int_0^\infty dx\ x \mid V(x) \mid < \infty, \tag{2.13a}$$

$$\int_0^\infty dx\ x\ e^{\alpha x} \mid V(x) \mid < \infty \quad \text{for any } \alpha > 0. \tag{2.13b}$$

Note that if condition (2.13b) is violated for some $\alpha > \alpha_o > 0$, then $g(\omega, x)$ may not be analytic for Im $\omega < -\alpha_o$. The singularities that appear for potentials that have a tail, e.g., inverse-power-law potentials, have been studied, with a focus on understanding the associated late time behavior [10].

It then remains to investigate the behavior of $\tilde{G}$ on the large semicircle.

B. Asymptotic behavior of Green's function

We first give a simple (and not strictly rigorous) derivation to highlight the essential ideas. The asymptotic behavior of $\tilde{G}$ at high frequencies can be determined using the WKB method. Let $\tilde{\phi}(x) \equiv \exp[iS(x)]$ be a solution of the time-independent KG equation; then in lowest order approximation

$$S(x) = \pm \int^x k(x')dx', \tag{2.14}$$

where the position-dependent wave number $k(x)$ is given by $k(x) = [\omega^2 - V(x)]^{1/2}$. The two auxiliary functions f and g, and hence the Green's function $\tilde{G}$, can be obtained in terms of $\tilde{\phi}(x)$ and $k(x)$.

Consider a potential with a discontinuity at $x = a$. In this situation the WKB approximation breaks down at the discontinuity; however, one can join the two approximate solutions across the discontinuity using the reflection coefficient

$$R = \frac{S'(a^-) - S'(a^+)}{[S'(a^-)]^* + S'(a^+)} \ . \tag{2.15}$$

The discontinuity can be in any derivative of V, and R behaves as some inverse power of ω as $|\omega| \to \infty$. (This will be spelt out precisely below.) We divide the discussion into two cases: (a) $0 < y < a < x$; (b) $0 < y \le x < a$. Referring to the discontinuity at $x = a$ as the stellar surface for simplicity, we may say that case (a) is relevant when gravitational waves from a source inside the star is detected by a distant observer outside the star, while case (b) is relevant for discussing the internal dynamics of the star.

For case (a), it is straightforward to show that

$$\tilde{G}(x,y;\omega) \simeq \frac{(1+R)\sin[I(0,y)]e^{iI(a,x)}}{\sqrt{k(x)k(y)}[e^{-iI(0,a)} + Re^{iI(0,a)}]} \ , \quad y < a < x, \tag{2.16}$$

where

$$I(u,v) = \int_u^v k(x)dx \approx \omega(v-u). \tag{2.17}$$

Now consider $\tilde{G}$ on the semicircle $\omega = \omega_R + i\omega_I = Ce^{i\theta}, \pi < \theta < 2\pi$, with $C \to \infty$. As $\omega_I \to -\infty$, both the numerator and the denominator of $\tilde{G}$ are dominated by the term proportional to R, and

$$\tilde{G}(x,y;\omega)e^{-i\omega t} \simeq \frac{(1+R)e^{-i\omega(t-x-y+2a)}}{R\omega}. \tag{2.18}$$

As $C \to \infty$, this vanishes for $t > t_p \equiv \max(x+y-2a, 0)$, since R varies as an inverse power of ω. Thus we have proved (modulo a technical complication to be mended below), that for a discontinuous potential, completeness in the sense (1.2) holds for $t > t_p(x,y)$.

Next consider case (b). Instead of (2.16), one now has

$$\tilde{G}(x,y;\omega) \simeq \frac{\sin[I(0,y)][e^{-iI(x,a)} + Re^{iI(x,a)}]}{\sqrt{k(x)k(y)}[e^{-iI(0,a)} + Re^{iI(0,a)}]} \ . \tag{2.19}$$

Again, both the numerator and the denominator in $\tilde{G}$ are dominated by the term proportional to R, and

$$\tilde{G}(x,y;\omega)e^{-i\omega t} \simeq \frac{e^{-i\omega(t+x-y)}}{\omega}, \qquad y \le x < a. \tag{2.20}$$

As $C \to \infty$, this vanishes for all $t > 0$. Thus completeness is again proved.

The above derivation shows the essence of the proof, which relies on the vanishing of an exponential factor $e^{-|\omega_I|\sigma}$ where $\sigma > 0$ (say $\sigma = t - x - y + 2a$), when $\omega = Ce^{i\theta}, \pi < \theta < 2\pi$ and $C \to \infty$. This derivation however must be mended in the domain $|\omega_I| = O(\log C)$, i.e., within an angle $\Delta\theta \sim \log C/C$ of the real axis, in which $e^{-|\omega_I|\sigma}$ behaves only as a power C^σ, rather than as an exponential [17]. The problem is closely related to the QNM's, which lie precisely in this domain.

To discuss this in a somewhat more general context, consider a potential with a discontinuity in $d^pV(x)/dx^p$ at $x = a$, for some $p \ge 0$. Then as $\omega \to \infty$, $R \sim A\omega^{-q}$, where $q = p + 2$.

We first determine the asymptotic position of the QNM's, which are given by the zeros of the denominator $\tilde{G}$ in (2.16):

$$e^{-iI(0,a)} + Re^{iI(0,a)} \simeq e^{-i\omega a} + A\omega^{-q}e^{i\omega a} = 0, \tag{2.21}$$

with the solution

$$\omega_j a \simeq j\pi - i\left[q\log(j\pi) - \log(-a^q A)\right]/2, \tag{2.22}$$

where j takes an integer values. The distribution of QNM's is shown schematically in Fig. 3.

The derivation sketched above relies on the fact that the integrand vanishes on the large semicircle as $C \to \infty$. However, for $\omega \simeq \omega_j$, one can readily obtain the following estimate for the integrand:

$$|\tilde{G}(x,y;\omega)e^{-i\omega t}| \sim \frac{|R|^{(t-x-y)/2a}}{|\omega|} \sim \frac{|\omega^{-q(t-x-y)/2a}|}{|\omega|}, \tag{2.23}$$

for both case (a) and case (b). In order to justify the derivation for the completeness of QNM, one could restrict $t > x + y$, but this would greatly limit the usefulness of the completeness relation. However, it will be shown in the next subsection that this difficulty can be surmounted by using a standard regularization scheme.

The situation for $t < x + y$ in case (b), which does not give convergence, is markedly different from the case of the wave equation with a step discontinuity, which was considered by us in the context of optics [11], and also by Price and Hunain as a model for gravitational waves [9]. In that case, the analog of the potential $V(x)$ is $\omega^2\epsilon(x)$, where $\epsilon(x) \equiv n(x)^2$ is the dielectric constant. Any discontinuity in $\epsilon(x)$ is therefore amplified as $\omega \to \infty$ by two powers; specifically, if there is a discontinuity in $d^p\epsilon(x)/dx^p$, then $R \sim A\omega^{-q}$ at high frequencies, with $q = p$. The difference by two powers also follows from the mapping (1.10). Thus for $p = 0$, $\tilde{G}$ is sufficiently bounded for the proof of completeness to go through [11], and the statements about the possibility of expressing G and $\partial G/\partial t$ as sums over QNM's are strictly valid for the wave equation [18]. If $p \geq 1$, then the same problem arises even for the wave equation.

However, both the case of the KG equation, and the case of the wave equation with $p \geq 1$, can be rectified by a simple scheme of regularization.

C. Regularization

First, it follows directly from the reflection symmetry relation $\tilde{G}(x, y; -\omega^*) = \tilde{G}(x, y; \omega)^*$ that

$$G(x, y; t) = 2\,\mathrm{Re} \int_0^\infty \frac{d\omega}{2\pi} e^{-i\omega t} \tilde{G}(x, y; \omega). \tag{2.24}$$

As the integrand is well behaved at both endpoints, we can multiply it by an extra factor $e^{-\omega\tau}$, $\tau > 0$, and then take the limit $\tau \to 0^+$ afterwards, i.e.,

$$G(x, y; t) = 2\,\mathrm{Re} \lim_{\tau \to 0^+} \int_0^\infty \frac{d\omega}{2\pi} e^{-i\omega(t - i\tau)} \tilde{G}(x, y; \omega). \tag{2.25}$$

This is equivalent to assigning t an infinitesimal negative imaginary part, and provides sufficient regulation on the part of the contour with $|\omega_I|$ small to make the integral vanish as the contour expands to infinity in the lower half plane.

Second, we deform the integration contour in (2.25) into a sequence of expanding rectangular contours $\Gamma_n = \Gamma_{n1} \bigcup \Gamma_{n2} \bigcup \Gamma_{n3}$ in the lower half ω plane as shown in Fig. 4. However, instead of letting the width of the contour Γ_n expand continuously to infinity (which would hit the QNM's), each Γ_n is defined to cut roughly mid-way between ω_n and ω_{n+1}, and eventually we take the limit $n \to \infty$. The height Ω_n of the contour is chosen to be proportional to n. As a consequence, G can be expressed as a sum over QNM's lying within the rectangle formed by the real axis and Γ_n [19], plus an integral along the contour Γ_n:

$$G(x, y; t) = 2\,\mathrm{Re} \int_{\Gamma_n} \frac{d\omega}{2\pi} e^{-i\omega(t - i\tau)} \tilde{G}(x, y; \omega) + \mathrm{Re} \lim_{\tau \to 0^+} \left[i \sum_{j>0}^{n} \frac{f_j(x) f_j(y) e^{-i\omega_j (t - i\tau)}}{\omega_j \ll f_j | f_j \gg} \right]. \tag{2.26}$$

Third, provided $V(x)$ does not have a tail, there are no cuts, and $\tilde{G}(x, y; \omega)$ is real along the imaginary axis, which also follows from the reflection symmetry of $\tilde{G}$. Thus the integral along Γ_{n1} is purely imaginary and does not contribute to G.

For the integral along Γ_{n2}, the original argument sketched in the previous subsection is valid: when Ω_n is sufficiently large, the integrand goes as $\exp(-\mid \omega_I \mid \sigma)$, $\sigma > 0$, and therefore vanishes in the limit $\Omega_n \rightarrow \infty$; the condition $t > t_p$ for case (a) is necessary in order that $\sigma > 0$.

Next, divide each contour Γ_{n3} into two parts, with $\mid \omega_I \mid$ large in one part, and $\mid \omega_I \mid$ small in the other. In the first part, the integral also vanishes as the contours expand to infinity, for the same reason mentioned previously. The complication, as indicated earlier, lies with the second part, where $\mid \omega_I \mid$ is not large, and therefore does not provide an exponential damping factor. However, in this part the integrand is at most a power of ω and the regulating factor $e^{-\omega\tau}$ causes it to vanish as the contours expand to infinity.

As the contribution from the contour Γ_n vanishes when $n \rightarrow \infty$, the Green's function $G(x, y; t)$ can be decomposed into a regulated sum of QNM's, namely

$$G(x,y;t) = \mathrm{Re} \lim_{\tau \rightarrow 0^+} \left[i \sum_{j>0} \frac{f_j(x) f_j(y) e^{-i\omega_j (t - i\tau)}}{\omega_j \ll f_j | f_j \gg} \right] . \tag{2.27}$$

Thus the prescription is, very simply, to (i) consider only $\omega_R > 0$ and (ii) where necessary give t a small negative imaginary part τ. (The exceptional case where some QNM's lie exactly on the imaginary axis is readily handled [19].)

The physical meaning of the regularization is as follows. The Green's function $G(x, y; t)$ is a well defined object, independent of the size of the contour Γ. We attempt to write it as the sum of a prompt part arising from the integral on a large semicircle (or other large contour linking $\omega = -i\infty$ to $\omega = +\infty$), and the sum over QNM's. When the QNM frequencies extend to infinity, and there are insufficient powers of $1/\omega$ in the integrand, each of the two contributions individually diverges as the contour expands to infinity [17]. Physically, in these circumstances, high frequency QNM's cannot be cleanly separated from the prompt contribution, which is hardly surprising. However, the device of letting $t \rightarrow t - i\tau$ effects such a clean separation which is practically useful, as illustrated in the numerical examples in Sec. IV.

Mathematically, the regulator $e^{-\omega_j \tau}$ can be replaced by any other regulator $I_j(\tau)$ satisfying $I_j(\tau) \to 1$ as $\tau \to 0$ and $I_j(\tau) \to 0$ as $j \to \infty$. One possible alternative is

$$\begin{aligned} I_j(\tau) &= (N - j + 1)/N, \quad j \le N, \\ &= 0, \quad j > N, \end{aligned} \tag{2.28}$$

where N is the least integer greater than $1/\tau$; this corresponds to defining the divergent series by a Cesaro sum [20].

In practice, one would use a small value of τ rather than take $\tau \to 0^+$. Consider, for example, the regulator $e^{-\omega_j \tau}$ and using G to propagate given initial data. Leaving τ finite incurs an error only for those modes with $|\omega_j| \gtrsim 1/\tau$, and hence only affects the resultant wavefunction on length scales below $\Delta \sim 1/|\omega_j| \lesssim \tau$. Thus, so long as only a finite spatial resolution is required, not taking $\tau \to 0^+$ has little effect, as demonstrated in the numerical example.

D. Full line problem

The proof of completeness can be generalized to include cases with multiple discontinuities and also to problems on a full line: $-\infty < x < \infty$. The latter is physically relevant if the background spacetime is due to a Schwarzschild black hole of mass M. Under the transformation (1.8), or more explicitly

$$x = r + 2M \ln(\frac{r}{2M} - 1), \tag{2.29}$$

the event horizon ($r = 2M$) maps into $x \to -\infty$, while $r \to +\infty$ maps to $x \to +\infty$. The main idea is still to construct the Green's function $\tilde{G}$ from the two homogeneous solutions of the KG equation and then examine its high frequency behavior using the WKB approximation. Consider a full line problem where there are n discontinuities in $V(x)$ at $x = a_1, a_2, ..., a_n$. We show only the case $a_1 \le y \le x \le a_n$; the case where the observation point is outside (a_1, a_n) can be demonstrated similarly. The QNM's are defined by (2.1) with the outgoing waves conditions at $x \to \pm\infty$. For simplicity we shall assume that $\lim_{x \to \pm\infty} V(x) = 0$, and consequently

$$\lim_{x\to\pm\infty} f_j(x) \sim \exp(i\omega \mid x \mid). \tag{2.30}$$

In addition, we assume that $V(x)$ has no tail on either side, in the sense of (2.13).

Introduce two auxiliary functions $g_-(\omega,x)$ and $g_+(\omega,x)$ which are the solutions to the time-independent KG equation $\tilde{D}g_\pm = 0$, with the boundary conditions $\lim_{x\to\pm\infty}[g_\pm(x)e^{\mp i\omega x}] = 1$. As before, the Green's function in frequency space is

$$\begin{aligned}\tilde{G}(x,y;\omega) &= g_+(\omega,x)g_-(\omega,y)/W(\omega) \quad \text{for } y<x,\\ &= g_+(\omega,y)g_-(\omega,x)/W(\omega) \quad \text{for } x<y,\end{aligned} \tag{2.31}$$

where the Wronksian $W(\omega) = g_+(\omega,x)g'_-(\omega,x) - g'_+(\omega,x)g_-(\omega,x)$ is again independent of x. Since $V(x)$ has no tail, $g_\pm(\omega,x)$ are analytic in ω, and hence $\tilde{G}$ is also analytic except at the zeros of $W(\omega)$. If one can show that the integral along the semicircle at infinity vanishes, the QNM's will form a complete set, as described in Sec. IIB.

Let us examine the asymptotic behavior of $g_\pm(\omega,x)$ along a large semicircle in the half plane using the WKB approximation. Assume that $a_1 < a_2 < ... < a_n$, then for $x > a_n$

$$g_+(\omega,x) \simeq \exp[iI(a_n,x)]g_+(\omega,x=a_n). \tag{2.32}$$

Now $g_+(\omega,x)$ in general will consist of two counter-propagating waves for $a_j \le x \le a_{j+1}$ and can be expressed as follows:

$$g_+(\omega,x) \simeq A_j\exp[iI(a_j,x)] + B_j\exp[-iI(a_j,x)]. \tag{2.33}$$

The coefficients A_j and B_j can be obtained recursively from the relation

$$\begin{aligned}\begin{pmatrix} A_j \\ B_j \end{pmatrix} &= \begin{pmatrix} M_{11}(j) & M_{12}(j) \\ M_{21}(j) & M_{22}(j) \end{pmatrix} \times \\ &\begin{pmatrix} \exp[iI(a_{j-1},a_j)] & 0 \\ 0 & \exp[-iI(a_{j-1},a_j)] \end{pmatrix} \begin{pmatrix} A_{j-1} \\ B_{j-1} \end{pmatrix}\end{aligned} \tag{2.34}$$

The boundary condition is $A_n = g_+(\omega, x=a_n)$ and $B_n = 0$. The transfer matrix $\mathbf{M}(j)$ joining the wavefunction across the discontinuity at $x = a_j$ can be obtained by matching the boundary conditions at the specified point. At high frequencies

$$M_{11}(j) = M_{22}(j) \simeq 1, \tag{2.35}$$

and

$$M_{12}(j) = M_{21}(j) \simeq -R_j, \tag{2.36}$$

where R_j is the reflection coefficient defined by (2.15) and evaluated at $x = a_j$.

The asymptotic form of $g_-(\omega, x)$ can be similarly expressed in terms of two counter-propagating waves for $a_j \leq x \leq a_{j+1}$,

$$g_-(\omega, x) \simeq C_j \exp[iI(a_j, x)] + D_j \exp[-iI(a_j, x)], \tag{2.37}$$

where the coefficients C_j and D_j obey the same recursion relation as A_j and B_j, but with the initial conditions

$$\begin{pmatrix} C_1 \\ D_1 \end{pmatrix} = \begin{pmatrix} M_{11}(1) & M_{12}(1) \\ M_{21}(1) & M_{22}(1) \end{pmatrix} \begin{pmatrix} 0 \\ g_-(\omega, x = a_1) \end{pmatrix} \tag{2.38}$$

Then for $x \leq a_1$

$$g_-(\omega, x) \simeq \exp[-iI(a_1, x)] g_-(\omega, x = a_1). \tag{2.39}$$

We now demonstrate the completeness of QNM's for $x \in (a_1, a_n)$ as follows. As $\omega \to \infty$ in the lower half plane, the two auxiliary functions can be simplified by keeping only waves which grow exponentially in their expansions, and consequently

$$g_+(\omega, x) \simeq -R_n \exp[-iI(a_n, x)] g_+(\omega, x = a_n), \tag{2.40}$$

$$g_-(\omega, x) \simeq -R_1 \exp[+iI(a_1, x)] g_-(\omega, x = a_1). \tag{2.41}$$

The Green's function is thus given by

$$\tilde{G}(x, y; \omega) \simeq \exp\{i[I(a_1, y) + I(x, a_n) - I(a_1, a_n)]\}/(2i\omega) \tag{2.42}$$

for $a_1 < y \leq x < a_n$, and vanishes as $\omega_I \to -\infty$. The completeness of QNM's then follows. There is the same problem in the domain $|\omega_I| = O(\log |\omega_R|)$, and the same regularization needs to be applied.

III. DIFFERENT SENSES OF COMPLETENESS

We are now in a position to specify precisely what we mean by completeness, and to summarize our results in terms of these definitions. Completeness can be expressed in different forms and the meanings may or may not be the same. The first sense ($C1$) shall mean the validity of the following expression, in a certain domain of x, y, t, for the Green's function G for problems with an outgoing boundary condition:

$$C1: \quad G(x,y;t) = \lim_{\tau\to 0} \mathrm{Re}\left[i \sum_{j>0} \frac{f_j(x) f_j(y) e^{-i\omega_j t} I_j(\tau)}{\omega \ll f_j | f_j \gg} \right]. \tag{3.1}$$

In (3.1), the source point is $(y, t' = 0)$ and the observation point is (x,t); for most cases of interest in gravitational systems, $y < x$. The sum is over all QNM's such that the complex frequency ω_j has Re $\omega_j > 0$; the eigenfunctions are $f_j(x)$ and $\ll f_j \mid f_j \gg$ is the generalized norm. The regulating factor I_j could be

$$I_j(\tau) = e^{-\omega_j \tau}, \tag{3.2}$$

which corresponds to evaluating $G(x,y;t)$ at a complex t with a small imaginary part.

We have shown that for discontinuous potentials without a tail, $C1$ holds if (a) $0 < y < a < x$ and $t > t_p \geq 0$; or (b) $0 < y$, $x < a$ and $t > 0$. We further note that with the regulating factor $I_j(\tau)$, for (x,t) in these two ranges (a) and (b), the sum (3.1) holds uniformly and it is valid to differentiate term by term. This leads to

$$C2: \quad \dot{G}(x,y;t) = \mathrm{Re}\left[\sum_{j>0} \frac{f_j(x) f_j(y) e^{-i\omega_j t} I_j(\tau)}{\ll f_j | f_j \gg} \right]. \tag{3.3}$$

$C2$ holds under the same conditions as $C1$, provided a suitable regulating factor is used.

The importance of $C1$ and $C2$ is that they determine the evolution of initial data. For $t > 0$, we have

$$\phi(x,t) = \int dy \left[G(x,y;t)\dot{\phi}(y,0) + \dot{G}(x,y;t)\phi(y,0) \right]. \tag{3.4}$$

The expansion of $G(x,y;t)$ and $\dot{G}(x,y;t)$ in terms of QNM's implies that the time development is uniquely determined by the QNM's. Therefore, for discontinuous potentials without tails, if $\phi(y,0)$ and $\dot{\phi}(y,0)$ have support only inside $(0,a)$, $\phi(x<a,t)$ is completely determined for all times by the QNM's, while $\phi(x>a,t)$ is completely determined by the QNM's for $t>t_p$.

How is this notion of completeness in terms of evolution related to the notion of completeness in terms of the expansion of the delta function? We note that from the defining equation and initial condition for G, $\dot{G}(x,y;t=0^+)=\delta(x-y)$; so if the domain of validity of $C2$ includes $t\to 0^+$, then one has in particular

$$C3: \quad \delta(x-y) = \mathrm{Re}\left[\sum_{j>0} \frac{f_j(x)f_j(y)I_j(\tau)}{\ll f_j|f_j \gg}\right], \tag{3.5}$$

which may be seen as the more familiar notion of completeness [18].

It is necessary to distinguish these notions of completeness because, unlike the familiar case of hermitian systems, completeness in the sense of evolution ($C1$ and $C2$) is not necessarily equivalent to completeness in terms of the expansion of the delta function ($C3$). This comes about because each of these holds only in *limited* regions of spacetime: If $y<a<x$, then $C1$ and $C2$ only hold for $t>t_p>0$; as t is bounded away from 0^+, $C3$ may not be valid. Conversely, suppose we know $C3$ to start with. In the familiar case of hermitian systems, such a resolution of the identity operator $\delta(x-y)$ would allow a decomposition of the initial data into a sum of eigenfunctions; attaching phase factors $\exp(-i\omega_j t)$ to each of these would then give the solution for dynamic evolution. This procedure gives an expression for the Green's function G for $t>0$, thus proving $C1$ and $C2$. This is why completeness in the sense $C3$ is usually regarded as the key concept for hermitian systems. However, in the present case, because $C3$ is valid only inside the "cavity" $(0,a)$, the superposition of eigenfunctions would only be valid if $x\in(0,a)$. Outside this interval, the initial condition would not be satisfied by a sum of eigenfunctions and consequently one does not have $C1$ and $C2$ from $C3$ in general.

There is another important case for which $C3$ does not lead to $C1$ and $C2$, which is useful to sketch here. Consider the case $V_o \equiv m_o^2 \equiv \lim_{x\to\infty} V(x) > 0$. In this case, the asymptotic wave number k and the frequency ω are related by $k=\sqrt{\omega^2-m_o^2}$. The asymptotic behavior would be, for example, $\lim_{x\to\infty}[g(\omega,x)\exp(-ikx)] = 1$. Most of the derivation is easily adapted, except that the relation between k and ω leads to an extra cut S on the real ω axis. The integral for $G(x,y;t)$ in terms of $\tilde{G}$ is above this cut, so in general, there will be this extra cut contribution when the contour is distorted to the lower half plane. Thus the QNM sum is *not* complete in the sense of $C1$ and $C2$. Nevertheless, it turns out that this cut contribution vanishes for $t \to 0^+$, so the QNM representation of $\delta(x-y)$, i.e., $C3$ remains valid.

To see this, define the integral around the cut S

$$\begin{aligned} G_S(x,y;t) &\equiv \int_S \frac{d\omega}{2\pi}\tilde{G}(x,y;\omega)\exp(-i\omega t), \\ &= \int_{-m_o}^{m_o}\frac{d\omega}{2\pi}[\tilde{G}(x,y;\omega+i0) - \tilde{G}(x,y;\omega-i0)]\exp(-i\omega t). \end{aligned} \tag{3.6}$$

As both $\tilde{G}(x,y;\omega+i0)$ and $\tilde{G}(x,y;\omega-i0)$ satisfy (2.2), their difference is evidently proportional to the product of the regular solutions $f(\omega,x)$ and $f(\omega,y)$. These functions diverge exponentially as x or $y \to \infty$ along the real axis, since the asymptotic wave number is now purely imaginary. However, they are normalizable along the imaginary x or y axis using the usual box normalization. It can be shown that G_S is expressible as an integral along the imaginary k axis as follows:

$$G_S(x,y;t) = \int_0^{im_o}\frac{dk}{\pi}\frac{\hat{f}(\omega,x)\hat{f}(\omega,y)\cos\omega t}{\omega}, \tag{3.7}$$

where $\hat{f}(\omega,x)$ is proportional to $f(\omega,x)$ and satisfies the box normalization condition

$$\lim_{L\to\infty}\frac{1}{iL}\int_0^{iL}\hat{f}(\omega,z)^2 dz = 1. \tag{3.8}$$

The frequency ω is now given by

$$\omega = \mid m_o^2 - k^2 \mid^{1/2}. \tag{3.9}$$

It is obvious from (3.7) that $\dot{G}_S(x,y;t=0^+)$ vanishes and does not contribute to the QNM decomposition of $\delta(x-y)$.

For open systems as studied here, the importance of $C3$ is twofold. For initial data having support in $(0, a)$, it enables the expansion of initial data in terms of QNM's. It is also important in the construction of a time-independent perturbation theory in terms of QNM's, an issue we investigated in the case of optical systems [11].

IV. NUMERICAL VERIFICATION

Consider a potential defined on the full line as follows

$$\begin{aligned} V(x) &= V_1 \quad \text{for} \;\; 0 \le x \le a, \\ &= 0 \quad \text{otherwise,} \end{aligned} \tag{4.1}$$

with $V_1 > 0$. This potential has a step discontinuity ($p = 0$) and no tail, thus satisfying the conditions necessary for our results. Moreover, such a repulsive potential has significant leakage, so it is particularly suited for verifying that our results are not merely valid to some low order in the amount of leakage. The QNM's of this system are determined by

$$\begin{aligned} f_j(x) &= \exp(i\omega x) \quad \text{for } x > a, \\ &= A\exp(ikx) + B\exp(-ikx) \quad \text{for } a > x > 0, \\ &= D\exp(-i\omega x) \quad \text{for } 0 > x, \end{aligned} \tag{4.2}$$

where $k = \sqrt{\omega^2 - V_1}$, and the QNM frequency ω_j is the j-th solution to

$$r(\omega)^2 \exp(2ika) = 1, \tag{4.3}$$

with reflection coefficient $r(\omega) \equiv (\omega - k)/(\omega + k) \sim 1/(4\omega^2)$ at high frequencies. The result (4.3) is a direct generalization of (2.21) by merely replacing $R(\omega)$ with $r(\omega)^2$, as waves are now reflected successively by the two discontinuities at $x = 0$ and $x = a$. The distribution of QNM's on the complex ω plane shown in Fig. 3 corresponds exactly to this model, and it is easy to show that

$$k_j a \simeq j\pi - 2i[\log(|j|\pi/(V_1^{1/2}a) + \log 2] \tag{4.4}$$

for $|j| \gg 1$, consistent with (2.22) with $R \sim \omega^{-4}$. Im ω increases with j, implying large leakage at high frequencies.

In the following we shall demonstrate numerically the following approximate equality for small τ

$$\mathrm{Re}\sum_{j>0}\frac{f_j(x)f_j(y)I_j(\tau)}{\ll f_j \mid f_j \gg} \simeq \delta(x-y), \tag{4.5}$$

where $I_j(\tau) = \exp(-\omega_j\tau)$. We have labeled the QNM's so that $j > 0$ corresponds to Re $\omega_j > 0$. Since (4.5) is expected to hold only in a distribution sense [18], we integrate y from y_1 to y_2

$$\mathrm{Re}\sum_{j>0}\frac{f_j(x)\left[\int_{y_1}^{y_2} dy f_j(y)\right] I_j(\tau)}{\ll f_j \mid f_j \gg} \simeq \theta(x-y_1) - \theta(x-y_2). \tag{4.6}$$

Denote the partial sum on the left up to $j = J$ as $S_J(x; y_1, y_2; \tau)$. Fig. 5a shows $|S_J|$ versus J for a case where $x \notin (y_1, y_2)$, and Fig. 5b shows $\mid S_J - 1 \mid$ versus J for $x \in (y_1, y_2)$. In each case, curves are shown for several values of τ. It is seen that for sufficiently small τ (in fact $\tau \ll \Delta = y_2 - y_1$), the equality indeed holds to a good approximation. These curves also show that convergence in j is more rapid for a slightly larger τ; this is natural from the regulating factor. This property implies that in practice, an extremely small τ may not be optimal.

To study the dependence on τ, Fig. 6a shows a function $\varphi_o(x)$ (e.g., some initial data) on the interval $(0, a)$, and Fig. 6b shows the absolute error in representing this function by the QNM sum, using $\tau = 10^{-2}$, 10^{-3}, 10^{-4}. It is seen that the error converges to zero as $\tau \to 0$, and that for most purposes, using a small finite τ does not matter in practice. The QNM sum does not converge if the regulator is removed.

It is useful to examine through this example why the series needs to be regulated from a broader point of view. Let $r(\omega) \sim \omega^{-q}$ at high frequencies ($q = 2$ for the present situation). By using a similar analysis as that applied in obtaining (2.22), it is straightforward to show that $\mathrm{Im}\, k_j a \sim -q \log |j|$. Next consider the middle region $0 < x < a$, and a typical term in the product $f_j(x)f_j(y)$

$$|f_j(x)f_j(y)| \sim e^{ik_j(x+y)} \sim (j\pi)^{q(|x/a-1/2|+|y/a-1/2|+1)}. \tag{4.7}$$

Moreover, it is readily shown that

$$\ll f_j \mid f_j \gg \sim (j\pi)^q. \tag{4.8}$$

Since the maximum value of $|x/a - 1/2| + |y/a - 1/2|$ is 1, the worst behavior of $|f_j(x)f_j(y)|/ \ll f_j \mid f_j \gg$ is $|j|^q$. It is seen that the sum (4.5) would not converge, even in a distribution sense, without a regulating factor if $q \geq 2$. The culprit is Im $k_j a \sim -q \log |j|$, from which one can say unequivocally that the need for regulating the sum is an intrinsic property of *open* KG systems.

It is worthwhile at this point to reiterate a difference between the wave equation (WE) and the KG equation. In each case, if there is a discontinuity in the p-th derivative (of the dielectric constant in one case and of the potential in the other case) then

$$\text{WE}: \quad q = p,$$

$$\text{KG}: \quad q = p + 2.$$

Thus in the case of the KG equation, the QNM sum requires regulation for all $p \geq 0$, whereas in the case of the WE, regulation is required only for $p \geq 2$. This is again natural when seen in the context of (1.10) — a step discontinuity in V is equivalent to a second-order discontinuity in the dielectric constant. This difference between the two systems, though of a rather technical nature, should nevertheless be kept in mind when using one system to draw conclusions about the other.

The need for a singularity is a feature that might seem, at first sight, to be surprising. Our results are valid for a discontinuity however soft (i.e., p however large), but are not valid for a C^∞ potential. Yet there should be little difference between a very soft discontinuity and no discontinuity. The resolution of this paradox lies in the regulator — for large p, the divergence is more severe, the regulated sum converges much more slowly to the correct result as $\tau \to 0$, and the use of a small finite τ becomes increasingly inaccurate. This behavior then connects smoothly to that for a C^∞ potential, for which the sum may not make sense for any τ.

V. DISCUSSIONS AND CONCLUSION

In this paper we studied the completeness of QNM's for linearized waves propagating in a curved background described by (1.1). It is well known that isolated Schwarzschild black holes do not have a complete set of QNM's, while other model systems described by (1.1) do have complete sets of quasinormal modes, e.g., the stellar oscillation model of [9]. The questions we investigate in this paper are: (i) What does completeness mean for the QNM's of an open system described by (1.1)? (ii) What properties of the potential would characterize those systems with complete sets of QNM's? The answers to these questions are spelt out in Secs. II and III.

We find that there are two important ingredients needed for completeness. First, the potential has to have no "tail" in the asymptotic region in the sense of (2.13). Furthermore, for potentials behaving asymptotically as $l(l+1)/x^2 + \bar{V}(x)$, where l is an integer, it can be shown that the completeness relation holds provided $\bar{V}(x)$ satisfies (2.13). Details of this generalization is given in other context [21]. On the other hand, if $\bar{V}(x)$ decays slower than an exponential, the QNM's may not form a complete set. It has been shown that for $\bar{V}(x) \sim 1/x^n$ with $n \geq 2$, the corresponding Green's function $\tilde{G}(\omega)$ has a branch cut along the negative imaginary ω axis, and consequently QNM's are incomplete [10]. The cut is related to the late time tail of gravitational waves [10], which decays only as an inverse power of time. This in turn verifies that QNM's do not form a complete set.

Second, the potential has to have discontinuities to provide a demarcation of a finite interval, analogous to the boundaries of an optical cavity. The discontinuities can be in any finite order of the spatial derivative of the potential. Discontinuities in the potential might seem unnatural, however we note that such discontinuities in V commonly exist in many models of gravitating systems, e.g., the Price potential as a model for the potential of a black hole [22], with the discontinuity representing the peak of the potential; and the stellar oscillation model of [9], with the discontinuity representing the stellar surface; as well as other stellar models [23,24].

Needless to say there are many models of gravitating systems for which these two conditions for completeness are not satisfied, with perhaps the most important example being an isolated Schwarzschild black hole. For these cases, we note that our study on completeness is still relevant in that (i) it provides understanding as to *why* the quasinormal modes of these systems should be incomplete, and (ii) it shows that their QNM's can be made complete by a small change in the potential, e.g., by setting the potential $V(x)$ to zero for $x > X$ for a large X where V is very small. On the one hand, one does not expect a small change to alter the physics much, yet on the other hand, the set of QNM's now becomes complete and can be a very useful tool in the analysis of black hole perturbation. This very interesting point deserves a more careful discussion, which we now provide.

Consider a smooth potential, labeled schematically as A; for example this could be the Schwarzschild potential. Approximate it by a discontinuous potential B, which vanishes rapidly at spatial infinity; this could, for example, be a piecewise constant approximation to a Schwarzschild potential [25]. Clearly the latter can be chosen such that the physics of the two systems are nearly the same; schematically PHY(A) $\simeq$ PHY(B). Yet it is not difficult to see, e.g., by reference to the sort of estimates such as (2.22), that the distribution of QNM's must be fundamentally different, i.e., QNM(A) $\neq$ QNM(B). Such a qualitative difference has been demonstrated numerically in the example of a piecewise constant approximation to Schwarzschild potential [25]. From our point of view, the important difference lies not only in the distribution of QNM's, but also in that QNM(A) is not complete, while QNM(B) *is* complete. These differences, both in the asymptotic QNM distribution and in the completeness, depend on the *order* of the discontinuity and not on the magnitude, and therefore does not go away even when B is very close to A. This situation is sometimes regarded as paradoxical [25]. The paradox is resolved if one realizes that the QNM's are not complete for system A, and their sum does not correctly describe the physics, i.e., PHY(A) $\neq$ QNM(A). Of course, the results in this paper implies PHY(B) = QNM(B). Thus, in all, we have QNM(A) $\neq$ PHY(A) $\simeq$ PHY(B) = QNM(B).

This remark opens up a particularly intriguing possibility. If we want to approximate the time development of A by a discrete sum of QNM's, we should *not* use the QNM's of the system itself. Instead, the QNM's of B, which are complete, can describe the evolution of A to a good approximation. Investigations along this line will be reported elsewhere.

ACKNOWLEDGMENTS

We thank H.M. Lai, K.L. Liu, S.Y. Liu, R. Price, and S.S. Tong for discussions. This work is supported in part by a grant from the Hong Kong Research Grant Council (Grant no. 452/95). The work of WMS is supported by the US NFS (Grant no. PHY 94-04788) and the CN Yang Fund while visiting The Chinese University of Hong Kong.

REFERENCES

[1] K.S. Thorne, Rev. Mod. Phys. **52**, 299 (1980).

[2] P.K. Kundu, Proc. R. Soc. London **A431**, 337 (1990); R.H. Price and J. Pullin, Phys. Rev. **D46**, 2497 (1992).

[3] It turns out that there is no such suppression, see R.H. Price, J. Pullin and P. K. Kundu, Phys. Rev. Lett. **70**, 1572 (1993).

[4] See, e.g., S. Chandrasekhar, *The Mathematical Theory of Black Holes*, Univ. of Chicago Press (1991).

[5] L. Lindblom and S. Detweiler, Astrophys. J. Suppl. **53**, 73 (1983), and references cited therein.

[6] C. V. Vishveshwara, Nature (London) **227**, 937 (1970); S. L. Detweiler and E. Szedenits, Astrophys. J. **231**, 211 (1979); L. Smarr, in *Sources of Gravitational Radiation*, ed. L. Smarr, Cambridge Univ. Press (1979); R. F. Stark and T. Piran, Phys. Rev. Lett. **55**, 891 (1985); P. Anninos, D. Hobill, E. Seidel, L. Smarr and W.-M. Suen, Phys. Rev. Lett. **71**, 2581 (1993); and references therein.

[7] See e.g., A. A. Abramovici et al., Science **256**, 325 (1992).

[8] E.S.C. Ching, P.T. Leung, W.M. Suen and K. Young, Phys. Rev. Lett. **74**, 4588 (1995).

[9] R.H. Price and V. Husain, Phys. Rev. Lett. **68**, 1973 (1992).

[10] E.S.C. Ching, P.T. Leung, W.M. Suen and K. Young, Phys. Rev. Lett. **74**, 2414 (1995); E.S.C. Ching, P.T. Leung, W.M. Suen and K. Young, "Wave Propagation in Gravitational Systems: Late Time Behavior", to appear in Phys. Rev. D.

[11] P.T. Leung, S.Y. Liu and K. Young, Phys. Rev. **A49**, 3057 (1994); P.T. Leung, S.Y. Liu, S.S. Tong and K. Young, Phys. Rev. A **49**, 3068 (1994); P.T. Leung, S.Y. Liu and K. Young, Phys. Rev. A **49**, 3982 (1994).

[12] We assume that here f' is nonzero at the origin. Otherwise, if $f \sim x^{l+1}$ as $x \to 0$, then we have $\lim_{x\to 0} x^{-(l+1)} f(\omega, x) = 1$.

[13] It is assumed that these are simple poles; generalization to higher order poles, or to cuts, is straightforward. For high order poles, the time dependence acquires an extra prefactor going like t^n; see e.g. J.S. Bell and C.J. Goebel, Phys. Rev. **B138**, 1198 (1965).

[14] Ya. B. Zeldovich, Zh. Eksp. Teor. Fiz. **39**, 776 (1960) [Sov. Phys. - JETP **12**, 542 (1961)].

[15] H.M. Lai, C.C. Lam, P.T. Leung and K. Young, J. Opt. Soc. Am. **B8**, 1962 (1991); P.T. Leung and K. Young, Phys. Rev. **A44**, 3152 (1991).

[16] R.G. Newton, J. Math. Phys. **1**, 319 (1960).

[17] Note that Jordan's Lemma states that an integral $\int d\omega e^{-i\omega t} f(\omega)$ along a large semicircle in the lower half plane will vanish for $t > 0$, provided that $|f(\omega)| \to 0$ at infinity. The necessary bound on f arises similarly from the need to control the part with $\mid \omega_I \mid = O(\log|\omega_R|)$.

[18] To be specific, in this paper we consider pointwise convergence for G (which is a well defined quantity) and convergence in the distribution sense for the representation of $\delta(x-y)$, which comes from $\dot{G}$. Going from G to $\dot{G}$ costs a factor ω, but going from the pointwise to the distribution sense effectively gains a factor ω^{-1}. This happens because asymptotically the typical dependence is $e^{i\omega y}$, so integration over y gives a factor ω^{-1}. Thus the pointwise validity of (2.12) and the validity of the QNM decomposition of δ-function in a distribution sense require the same conditions.

[19] If there are QNM's on the negative imaginary ω axis, one has to avoid going over these poles by taking the principal part of the integral in (2.26). As a consequence, the QNM sum of G in (2.12) and (2.27) should also include *half* of the corresponding sum over those QNM's with Re $\omega_j = 0$.

[20] G.H. Hardy, *Divergent Series*, Oxford, Clarendon Press (1948).

[21] P.T. Leung and K.M. Pang, "Completeness and Time-independent Perturbation of Morphology-dependent Resonances in Dielectric Spheres", preprint, Chinese University of Hong Kong (1995).

[22] S. Chandrasekhar and S. Detweiler, Proc. R. Soc. Lond. A. **344**, 441 (1975).

[23] K.D. Kokkotas and B.F. Schutz, Mon. Not. R. Astron. Soc. **255**, 119 (1992).

[24] K.D. Kokkotas, Mon. Not. R. Astron. Soc. **268**, 1015 (1994).

[25] H.-P. Nollert, "Quasinormal Frequencies of Step potentials", preprint, Universität Tübingen (1994).

FIGURE CAPTIONS

Fig. 1. The integrand in (2.10) is defined by the contour $L = L_1 + L_2$, where $L_1 = (0, a)$ and L_2 is the contour shown in this figure.

Fig. 2. Contributions from the poles, the semicircle, and the singularities of f and g (which in general form a cut on the negative imaginary axis).

Fig. 3. Distribution of QNM's for the model in Sec. IV with $a = 1$ and $V_1 = 100$.

Fig. 4. Contributions from the poles, and the rectangular contour $\Gamma_n = \Gamma_{n1} \bigcup \Gamma_{n2} \bigcup \Gamma_{n3}$.

Fig. 5. (a) Plot of log $|S_J|$ vs J for a case where $x \notin (y_1, y_2)$. (b) Plot of log $|S_J - 1|$ vs J for a case where $x \in (y_1, y_2)$. In each case $a = 1$, $V_1 = 100$, J is in thousands, and S_J is the partial sum in (4.6), evaluated for $\tau = 5 \times 10^{-3}$, 5×10^{-4}, 2×10^{-4} for lines 1, 2, 3 respectively. The partial sums are in fact fluctuating functions of J, and the lines shown are smooth envelopes representing upper bounds. These curves show that (i) at fixed τ, the sum over j converges; (ii) the resultant error vanishes as $\tau \to 0$.

Fig. 6. (a) A function $\varphi_o(x) = \sin(\pi x/a)$ defined on $(0, a)$. (b) The absolute error in representing this function by the QNM sum, using $\tau = 10^{-2}$ (circles), 10^{-3} (squares), 10^{-4} (triangles). In this example, $a = 1$ and $V_1 = 100$.