

PAPER

# Sensitivity kernels in seismic wave propagation: a simplified explanation for the banana-doughnut paradox

To cite this article: Caio Ciardelli 2022 *Eur. J. Phys.* **43** 045802

View the [article online](#) for updates and enhancements.

## You may also like

- [Spectral noise for edge states at the filling factor = 5/2](#)  
M Carrega, D Ferraro, A Braggio et al.
- [Finite-frequency Sensitivity Kernels in Spherical Geometry for Time-Distance Helioseismology](#)  
Krishnendu Mandal, Jishnu Bhattacharya, Samrat Halder et al.
- [Submicron hollow spot generation by solid immersion lens and structured illumination](#)  
M-S Kim, A C Assafrao, T Scharf et al.



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Sensitivity kernels in seismic wave propagation: a simplified explanation for the banana-doughnut paradox

Caio Ciardelli\* 

Department of Geophysics, Institute of Astronomy, Geophysics, and Atmospheric Sciences, University of São Paulo, Brazil

E-mail: [caio.ciardelli@gmail.com](mailto:caio.ciardelli@gmail.com)

Received 9 November 2021, revised 4 March 2022

Accepted for publication 26 April 2022

Published 9 June 2022



CrossMark

## Abstract

Ray theory, a high-frequency approximation to describe wave propagation, has been a cornerstone in seismology for over a hundred years. Despite its simplicity and wide range of applications, some limitations combined with the ever-increasing computational power motivated the development of finite-frequency theory, a better model to describe how the Earth's inner structure affects seismic waves. Finite-frequency theory has matured a lot in the last decades, and it is now widely applied in many geophysical problems. However, most students and even some experienced researchers face difficulties understanding it. An appropriate theoretical comprehension is paramount to making the most out of the methods a theory underpins, avoiding pushing it beyond its limits, and further developing it. With that problem in mind, this paper shows a simplified formulation of the sensitivity kernels, which are the generalization of rays in the finite-frequency regime. The resultant model, despite its limitations, correctly predicts the main features of finite-frequency theory, including the zero sensitivity in the middle of the travel-time kernels, known as the *banana-doughnut paradox*, shedding new light on that intriguing phenomenon. The step-by-step derivation and relatively easy equations should be understandable by an undergraduate student with a reasonable knowledge of classical physics and calculus. A Colab Notebook implementing the main formulas accompanies the paper, allowing readers to interact and play with the results.

Keywords: finite-frequency theory, sensitivity kernels, banana-doughnut paradox

(Some figures may appear in colour only in the online journal)

\*Author to whom any correspondence should be addressed.

## 1. Introduction

The increasing levels of specialization and automation are two crucial factors that have allowed both the ever-growing accumulation of knowledge and the technological progress that have brought countless benefits to humankind. In the last decades, science has experienced remarkable theoretical, experimental, and computational advances, and seismology is no exception. That enormous progress brought about the development of a wide range of methods, computational codes, and instruments that are now available to scientists. The development of high-level computational languages and the availability of many sophisticated computational packages made many activities very easy and productive. Despite all the advances and undeniable benefits of computational developments and automation, many scientists fail to make the most out of them because they do not have a deep understanding of the methodologies, techniques, and codes they are using. In geosciences, for example, one could mention the example of seismic tomography, a technique that uses seismic waves to image the Earth's interior. A variety of phenomena can create seismic waves, including earthquakes, asteroid impacts, and artificial explosions. Large earthquakes are the main source of seismic data for carrying out seismic tomography on a global scale. Provided that one knows the location and origin time of the earthquakes and that a network of seismographic stations is available, it is possible to use the record of the seismic waves at the receivers to infer the composition of the Earth's interior. The more earthquakes and stations spread over the surface of the planet, the better the result.

Historically, seismologists have been carrying out that process by using a velocity model, which is a simplified mathematical representation of the Earth. Using the velocity model, one computes the theoretical travel times, amplitudes, or even waveforms of the seismic waves and compares them with the observed values recorded at the stations. If the predicted values match the observed ones, we know our model is a reasonable representation of the planet's interior. Otherwise, we iteratively update the model until the difference between them is minimized subject to some physical, geological, and mineralogical constraints (e.g. the seismic velocities should lie in the range of known values for the Earth's rocks, derived from lab experiments with samples, and the model should not have excessive roughness).

Aki and Lee (1976) and Aki *et al* (1977) used teleseismic  $P$ -wave travel-time residuals (differences between observed and theoretical travel times) to determine the three-dimensional seismic structure of the lithosphere. Sengupta and Toksöz (1977) computed a 3D velocity model for the whole mantle combining  $P$ ,  $S$ , and some  $PcP$  ( $P$  wave reflected on the outer core) and  $ScS$  ( $S$  wave reflected on the outer core) travel-time measurements. Dziewonski *et al* (1977) imaged the top 1,100 km of the mantle using  $P$ -wave residuals. Since then, many 1D reference Earth models (e.g. Dziewonski and Anderson 1981, Kennet 1991, Kennett *et al* 1995) and 3D Earth models (e.g. Ritsema *et al* 1999, Grand 2002, Montelli *et al* 2006, Houser *et al* 2008, Ritsema *et al* 2011, Schaeffer and Lebedev 2013, French and Romanowicz 2014, Koelemeijer *et al* 2016, Durand *et al* 2017, Lu *et al* 2019, Lei *et al* 2020) using multiple techniques and various types of data have been developed, such as body-wave travel times, surface-wave dispersion, Earth's normal modes, mass and moment of inertia, and waveform inversion.

In the early years of seismic tomography, the limited computational power required low-cost computational methods that used seismic ray theory. The seismic ray theory (or simply 'ray theory') is an extremely important method in seismology, consisting of a high-frequency approximation for the wave field, borrowed from optics, in which elementary waves, such as  $P$ ,  $S$ , reflected, refracted, and transmitted waves, propagate along rays that can be handled independently (Cerveny 2005). Ray theory allows for easy and efficient computation of travel times and amplitudes of seismic waves. However, it only works properly in piecewise smooth

media, in which the dominant wavelength is considerably smaller than the dimensions of the inhomogeneities (Arora *et al* 2011). It also fails to take scattering effects into account. The limitations of the ray theory led to the development of the so-called finite-frequency theory in seismology. In that new theory, the high-frequency approximation is relaxed, and volumetric structures named *Fresnel volumes* (Brokesčová 2006), *Fréchet kernels* (Tarantola 1987), or *sensitivity kernels* (Marquering *et al* 1998) replace the infinitesimal seismic rays. The term ‘sensitivity’, in this case, refers to the fact that these structures describe what waves ‘sense’ in the Earth’s interior as they travel through it.

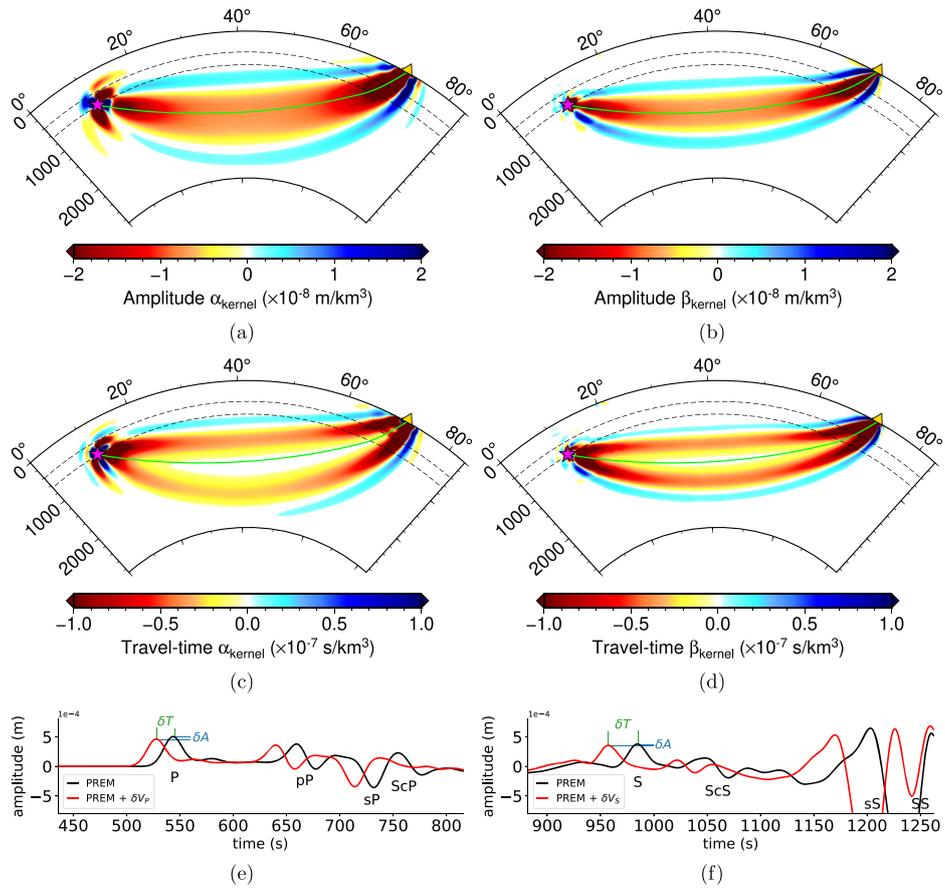
This paper presents a simplified derivation of these sensitivity kernels, which helps students of seismology come to terms with some difficult concepts of the finite-frequency theory, such as the so-called *banana-doughnut paradox*. Section 2 introduces the sensitivity kernels and the main questions regarding the finite-frequency theory. Section 3 briefly reviews Huygens’ principle and Fermat’s principle of wave propagation. Section 4 contains the derivation of the simplified model and its interpretation, explaining the absence of sensitivity in the middle of the travel-time kernels. Section 5 presents a Colab notebook that implements the main formulas, allowing the readers to interact with the model. Section 6 further discusses the results, and section 7 concludes the paper by summarizing the main takeaways.

## 2. Sensitivity kernels

Whenever one uses ray theory in seismic tomography, the implicit assumption is that any velocity anomaly not crossed by the ray does not affect the amplitude or the travel time of the recorded signal, since all the energy is assumed to travel along the ray. Another way of saying this is that all the amplitude and the travel-time sensitivities lie along the ray path. The fact that the sensitivity kernels have a non-zero volume in the finite-frequency regime shows that structures outside the ray path can affect the seismograms. The sensitivity kernels have a volume proportional to the prevailing wavelength of the seismic waves. As one would expect, the shorter the wavelength, the thinner the kernels. As the wavelength approaches zero, they collapse into geometrical rays with infinitesimal width, recovering ray theory. According to Snieder (1999) and Hung *et al* (2001), the kernels’ width scale with  $\sqrt{\lambda L}$ , where  $\lambda$  is the wavelength and  $L$  is the distance between source and receiver.

Besides the wavelength, the kernels also depend on the kind of wave ( $P$ ,  $S$ ,  $Pp$ ,  $ScS$ , *Rayleigh*, *Love*, etc) and the measurement. Amplitude measurements generate volumetric kernels whose shape resembles a banana. The maximum sensitivity occurs in the middle of the volume, where ray theory predicts the ray path. However, as first realized by Woodward (1992) and later by Marquering *et al* (1999), travel-time measurements give origin to what seems to be a paradoxical structure: a hollow banana. That result, at first sight, seems strange because it implies that right at the ray path, where one would expect the sensitivity to be maximum, the influence on the travel time of the waves is zero. Following Marquering *et al* (1999), Bozdağ *et al* (2011), and using SPECSEM3D\_GLOBE (Komatitsch and Tromp 2002a, 2002b), we can compute very accurate amplitude and travel-time kernels (figure 1).

The sensitivity kernels are also known as Fréchet kernels because they are a kind of Fréchet derivative. They show how the amplitude or the travel time varies with respect to an infinitesimal perturbation in a model parameter. We see in figure 1 that all the kernels are predominantly negative. This is because increasing the propagation velocity of the seismic waves has the double effect of reducing both the amplitudes and the travel times (figures 1(e) and (f)). The reduction in the travel time is easy to understand, but the effect on the amplitude is less obvious. To understand this, let us imagine a wave on a violin string, in which the propagation velocity



**Figure 1.** (a)–(d) Amplitude and travel-time sensitivity kernels for  $P$  and  $S$  waves. The amplitude kernels connecting the source (magenta star) and the receiver (yellow triangle) are solid, while the travel-time kernels are hollow, with zero sensitivity at the ray path (green line). In both cases, the dominant period is  $\sim 37$  s. The first and the second Fresnel zones are clear in all figures. The  $S$ -wave kernels ( $\beta$ ) are narrower than the  $P$ -wave kernels ( $\alpha$ ) because the  $P$ -wave velocity is approximately  $\sqrt{3}$  times larger than the  $S$ -wave velocity, resulting in a shorter wavelength. The velocity model is the isotropic version of the preliminary reference Earth model PREM (Dziewonski and Anderson 1981) and the black dashed lines denote the 410 and 660 km seismic velocity discontinuities. (e)  $P$  wave for unperturbed PREM (black) and perturbed with a 3% increase in the  $V_P$  (red). This velocity increase perturbs both the amplitude ( $\delta A$ ) and the travel time ( $\delta T$ ); (f) same as (e) but for an  $S$  wave before and after a 3% increase in the  $V_S$ .

increases with the string tension. The amount of work the violinist does on the string is proportional to both the amplitude of the deformation and the string tension. Assuming that the violinist always pulls the string with the same force, the amount of deformation is inversely proportional to its tension. Nevertheless, the energy of the resultant wave remains constant (a larger displacement is compensated by a weaker force). Therefore, for waves with the same energy content, a higher-tension string means higher velocity and smaller amplitude at the same time. For seismic waves, increasing the bulk modulus (resistance to volume changes) or

the shear modulus (rigidity) of the medium increases the velocity and decreases the amplitude for the same reason.

Ishimaru (1978) and Aki and Richards (1980) discuss the effects of scattering in wave propagation and show cross-sections of the sensitivity kernels, derived from Fresnel's wave theory. Groenenboom and Snieder (1995) show that the properties of the direct wave are determined by a weighted average over the first Fresnel zone for strongly scattering media. That result, in combination with the fact that in practical tomographic inversions, one always imposes some level of smoothness through model regularization, explains why in some cases ray theory gives correct results even for media where short-length-scale perturbations (smaller than the prevailing wavelength) are present (Snieder and Lomax 1996). Marquering *et al* (1999) and Hung *et al* (2001) explain why the travel-time kernels are hollow using scattering and a finite-frequency phenomenon known as *wavefront healing*. Spetzler and Snieder (2004) make an intuitive review of the finite-frequency theory, remarking on the limitations of ray theory and showing the connection between the sensitivity kernels and the Kirchhoff integral. The authors highlight that the advantage of the sensitivity kernels is that the integration is performed over a volume in the medium, whereas the Kirchhoff integral is defined on a surface only. Nolet *et al* (2005) present a simplified derivation of the amplitude and travel-time sensitivity of body waves with respect to velocity perturbations in the Earth's structure, taking into account the effect of finite frequencies. Nolet (2008) covers the essential aspects of seismic tomography, including the finite-frequency theory, at a level accessible to students.

Despite these excellent references, many phenomena related to finite-frequency effects remain mysterious for many people working in the field. As it might be the case that navigating through a more accessible approach could lead to a better understanding of the described phenomenon, the creation of a kernel model that is both simple and representative of its main features could help shed some light on the topic. Developing and exploring such a model is the subject of the rest of this paper.

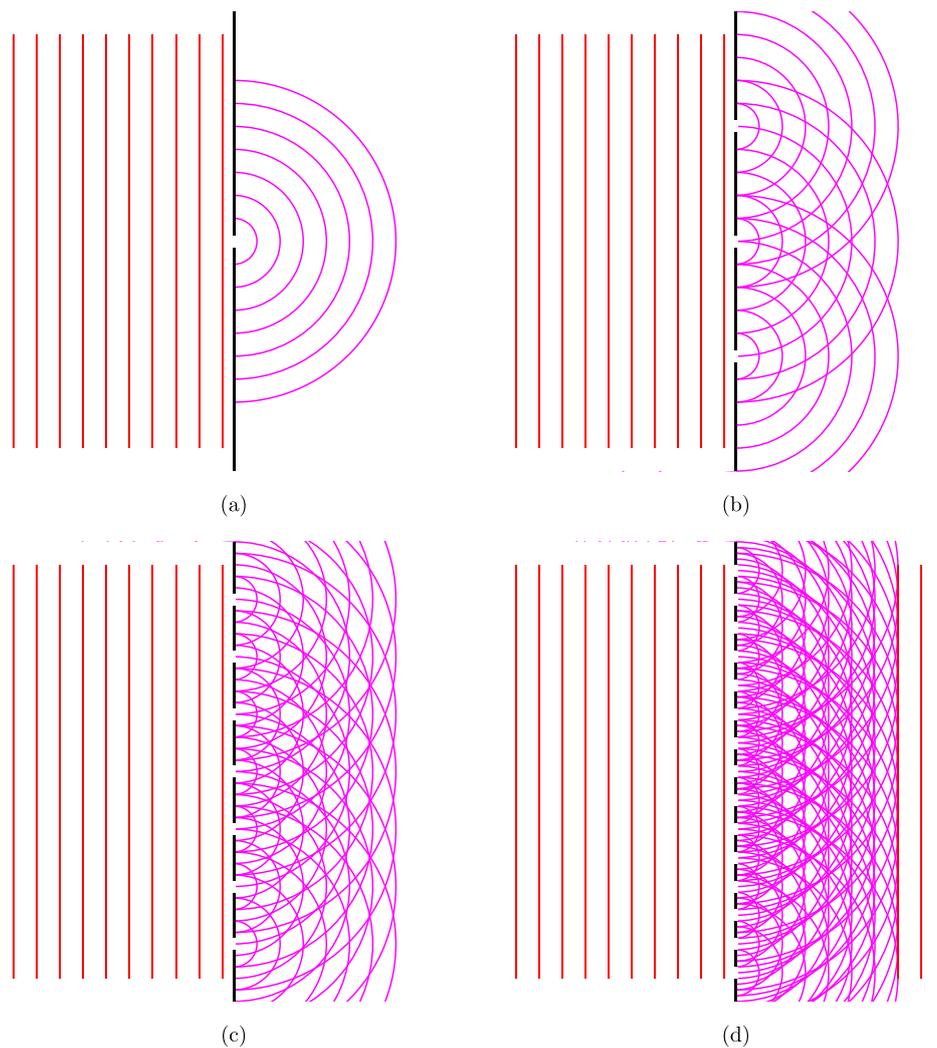
### 3. Basic principles of wave propagation

To create a model, it is convenient to review two principles that apply to undulatory phenomena: Huygens' principle (also known as Huygens–Fresnel principle) and Fermat's principle.

#### 3.1. Huygens' principle

Huygens' principle states that every point on a wavefront is also a source of spherical wavelets that propagate in all directions. The result of the superposition and interference of all these wavelets is the wavefront itself. An intriguing question concerning the Huygens' principle then arises: *if every point on the wavefront creates a new wave that propagates in all directions, then why does each wavefront not create two new wavefronts: one moving outwards, away from the source, and another traveling inwards, toward it? Why do we only observe the outward propagating wavefront?* Huygens' principle itself does not explain this, but one can demonstrate that only the outward propagating waves can exist by using the laws of conservation of energy and momentum.

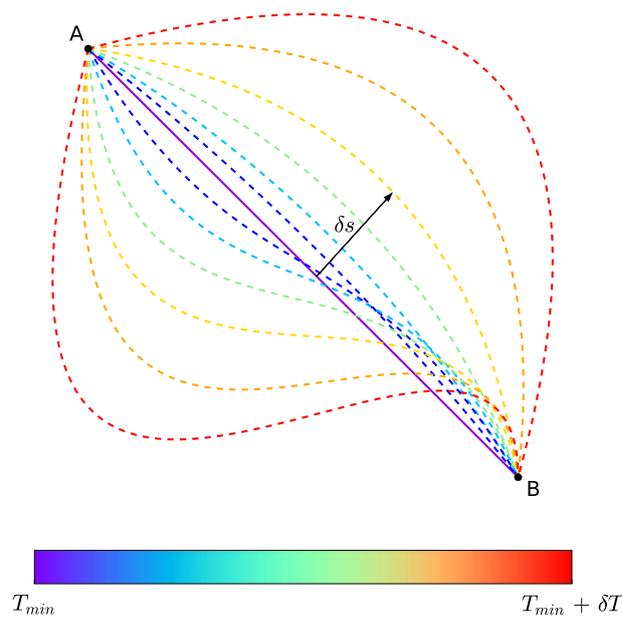
As mentioned by Robinson and Clark (2017), that problem was later solved by Fresnel and Kirchhoff with a more advanced formulation of Huygens' principle, expressed in the Fresnel–Kirchhoff integral theorem. Explaining that theorem is beyond the scope of this paper. However, in short, it conveys that the amplitudes of the spherical wavelets vary with the direction so that the amplitude of the resultant inward wavefront becomes zero. An equivalent way of stating this is that the inward propagating wavelets cancel each other out, resulting in a null inward wavefront.



**Figure 2.** Cross-section of a diffraction experiment using: (a) one slit; (b) three slits; (c) seven slits; and (d) fifteen slits. As the number of slits tends to infinity, the interference pattern recovers the plane wave.

**3.1.1. Huygens' principle and diffraction.** An easy way to visualize Huygens' principle is by observing diffraction phenomena. Suppose a plane wave in a 3D space is propagating from left to right and, eventually, the wavefront hits a wall with a single slit (figure 2). The wall blocks the waves created on almost all points on the wavefront, except those located right at the slit position. With a single slit, only a spherical wavefront emerges on the other side of the wall. However, what happens when we add more slits to the experiment? With two or more slits, an interference pattern emerges. If we add many slits, the waves interfere so that the resultant wavefront resembles the plane wave. With infinitely many slits, which is the equivalent of having no wall, we perfectly recover the plane wavefront.

A remark about Huygens' principle is that it does not hold when the number of spatial dimensions is even. In that case, the Green's function that solves the corresponding wave



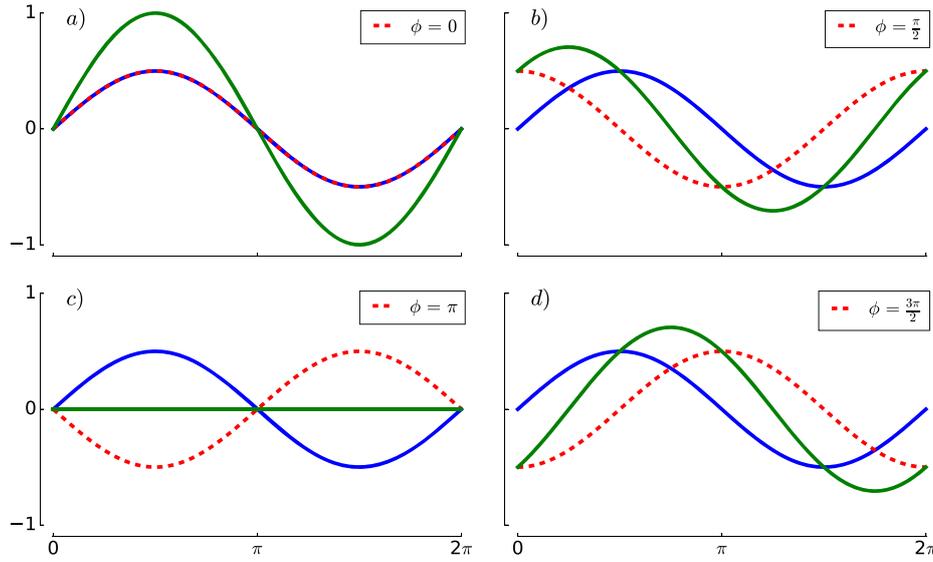
**Figure 3.** Illustration of Fermat's principle. In a homogeneous medium, the path of the least time for a wave created at point **A** and detected by a receiver at point **B** is a straight line (purple line). All the trajectories that are close to the least time trajectory (blue lines) have travel times very similar to the minimum ( $T_{\min}$ ). As the perturbation  $\delta s$  on the path gets larger (red lines), the travel time increases by a progressively larger  $\delta T$ .

equation has a tail, which implies that the information on the wavefront is not enough to predict its future state. Under those circumstances, one must also consider all the information in the space between the source and the wavefront (Dai and Stojkovic 2013).

### 3.2. Fermat's principle

According to Fermat's principle, a beam of light traveling from point **A** to point **B** always takes a path that is stationary with respect to small perturbations in the trajectory (figure 3). This principle is also commonly referred to as the principle of least time because the path of minimum time is the most common kind of stationary path. Nevertheless, maximum time paths or any other kind of stationary path are also possible according to Fermat's principle. In seismology, one could mention the *PP* (*P* wave bouncing off the surface once), *SS* (*S* wave bouncing off the surface once), *PS* (*P* wave bouncing off the surface once and being converted to an *S* wave), and *SP* (*S* wave bouncing off the surface once and being converted to a *P* wave) elastic waves as examples. These waves are called minmax phases because the partial derivatives of the travel time with respect to changes in the position of the bouncing point on the surface along the source-receiver direction and the orthogonal direction are both zero. However, the bouncing point is a saddle point of the travel time function: maximum in the source-receiver direction and minimum in the orthogonal direction.

However, how do waves 'know' which one is the path of the least time? Feynman (2006) explains why a photon bouncing off a mirror reflects with identical incidence and emergence angles (path of least time). He does so by using the quantum description of light, which allows the photon to simultaneously take all possible paths from the source to the receiver. He also



**Figure 4.** The interference can be constructive or destructive. (a) When the phase shift ( $\phi$ ) between the two signals (blue line and dashed red line) is zero, there is complete construction (green line). (b) When  $\phi = \frac{\pi}{2}$ , partial construction occurs. (c) When  $\phi = \pi$ , there is complete destruction. (d) When  $\phi = \frac{3\pi}{2}$ , partial construction occurs again.

shows that only the paths next to the one of the least time contribute to the resultant recorded signal. The key to understanding this is to realize that the travel time of a stationary path is invariant under small perturbations in the trajectory (figure 3). As a result, all paths close to that of the least time arrive at the receiver in phase and constructively interfere. Because all the remaining paths result in waves with randomly distributed phases, they cancel each other out (figure 4). This provides some insight regarding the dependence of the kernels' width on the wavelength. The larger the wavelength, the more the travel time can change before the waves get out of phase and destructively interfere. In the infinite-frequency regime, the wavelength is zero. In this case, a tiny difference in the travel time is enough to misalign the waves so that only the stationary path survives, recovering the ray theory.

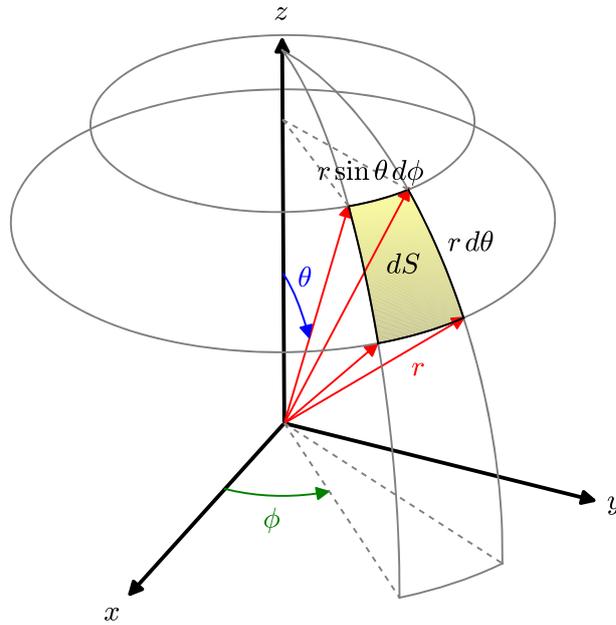
#### 4. Model formulation

We begin our simplified model of a sensitivity kernel by defining a convenient spherical coordinates system, shown in figure 5. A surface element in this coordinate system is given by:

$$dS = r^2 \sin \theta d\theta d\phi. \quad (1)$$

We calculate the surface area of a sphere  $S_S$  with radius  $r$  by integrating over the domain  $S = \{0 \leq \theta \leq \pi \text{ and } 0 \leq \phi \leq 2\pi\}$ :

$$S_S = \int_S dS = \int_0^{2\pi} \int_0^\pi r^2 \sin \theta d\theta d\phi = r^2 \int_0^{2\pi} d\phi (-\cos \theta|_0^\pi) = 4\pi r^2. \quad (2)$$



**Figure 5.** System of spherical coordinates in which  $r$  is the radius (red arrows),  $\theta$  is the colatitude (blue curved arrow), and  $\phi$  in the longitude (green curved arrow). The yellow region represents an infinitesimal element of area  $dS$ .

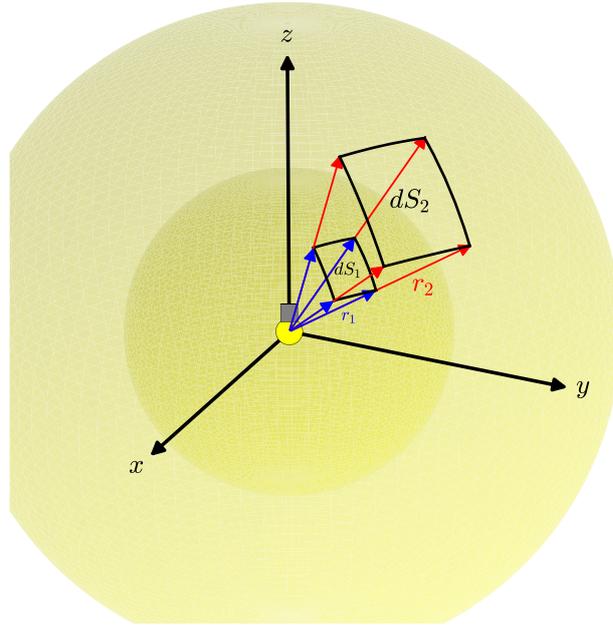
#### 4.1. Inverse-square law

As a next step, we place a light source at the origin of the coordinate system. The total energy output per unit time is  $E$ . We define the energy density on the wavefront ( $\varepsilon_S$ ) by the energy output divided by the area of the wavefront, which is the surface of a sphere. From that result, it is easy to derive the inverse-square law. The inverse-square law says that the energy density on the wavefront in a three-dimensional space is inversely proportional to the distance from the source squared, as shown in equation (3) and figure 6:

$$\varepsilon_S(r) = \frac{E}{4\pi r^2} \quad \text{and} \quad \frac{\varepsilon_{S2}}{\varepsilon_{S1}} = \frac{dS_1}{dS_2} = \frac{r_1^2}{r_2^2}. \quad (3)$$

In case we want the energy contained in the spherical wavefront ( $E_S$ ), we just need to integrate the energy density element  $dE = \varepsilon dS$  over the surface  $S$  (equation (4)). The result, as expected, is exactly the total energy output  $E$  of the source:

$$\begin{aligned} E_S &= \int_S dE = \int_S \varepsilon_S(r) dS = \int_0^{2\pi} \int_0^\pi \varepsilon_S(r) r^2 \sin \theta d\theta d\phi \\ &= \frac{E}{4\pi r^2} \int_0^{2\pi} \int_0^\pi r^2 \sin \theta d\theta d\phi \\ &= \frac{E}{4\pi r^2} \cancel{4\pi r^2} = E. \end{aligned} \quad (4)$$



**Figure 6.** We represent the energy source as a light bulb at the origin of the coordinate system. The energy density ratio between the two wavefronts is equal to the inverse ratio of their radii squared, as derived in equation (3).

#### 4.2. Energy incident on a virtual plane

We proceed by setting a virtual plane at a distance  $r$  from the energy source, and defining a new coordinate system, as depicted in figure 7. This scheme is similar to the parametrization used by Tian *et al* (2007). In this case, how much energy per unity time does the plane receive? In other words, if that plane was a 100% efficient solar panel, how much energy would it generate?

Using the parametrization described in figure 7, we calculate the total energy reaching the virtual plane by integrating all the light rings from  $s = 0$  to infinity. In this new coordinate system, the surface element  $dS$  is given by:

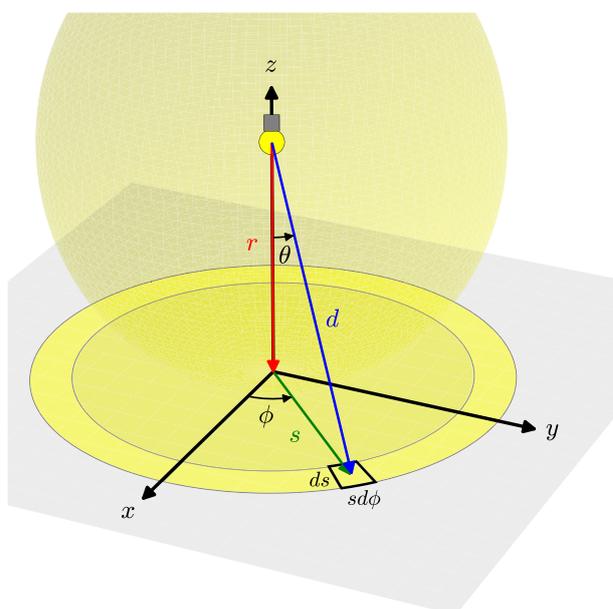
$$dS = s ds d\phi. \tag{5}$$

The first thing we need is the energy density on each ring ( $\epsilon_D$ ). Unlike the previous case, the energy density on the surface of the plane is not constant. It not only decays with the square of the distance  $d$  but also depends on the incidence angle  $\theta$  (Lambert's cosine law). Just as sunlight during winter spreads over a larger area due to a higher angle of incidence, the energy density on the disc decays with the cosine of  $\theta$ . Thus, combining the geometrical spreading with the effect of the oblique incidence, we find:

$$\epsilon_D(d) = \frac{E}{4\pi d^2} \cos \theta, \quad \text{where } d = \sqrt{r^2 + s^2} \text{ and } \cos \theta = \frac{r}{\sqrt{r^2 + s^2}}. \tag{6}$$

We rewrite equation (6) in terms of  $r$  and  $s$ :

$$\epsilon_D(r, s) = \frac{E}{4\pi(r^2 + s^2)} \frac{r}{\sqrt{r^2 + s^2}}. \tag{7}$$



**Figure 7.** Schematic describing a plane located at a distance  $r$  from the light source. The energy first arrives at the closest point on the plane, corresponding to  $s = 0$ . As time progresses, more and more light reaches the plane as a series of concentric rings of width  $ds$ .

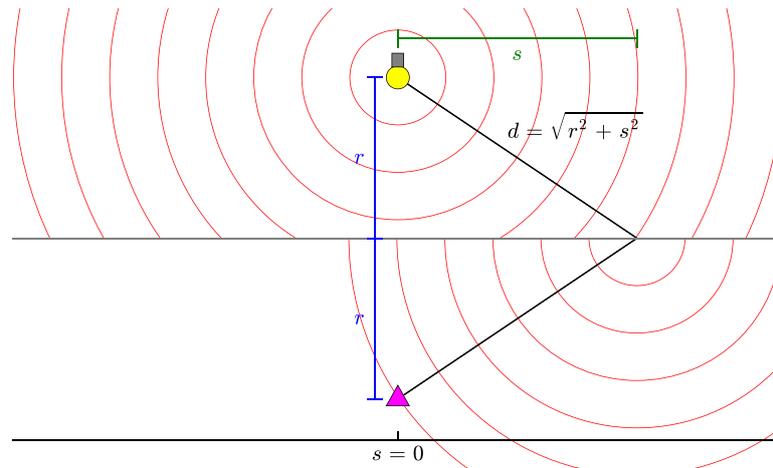
Therefore, the total energy reaching the plane  $E_D$  is:

$$\begin{aligned}
 E_D &= \int_S dE = \int_S \varepsilon_D(r, s) dS = \int_0^{2\pi} \int_0^\infty \varepsilon_D(r, s) s \, ds \, d\phi \\
 &= \int_0^{2\pi} \int_0^\infty \frac{E}{4\pi(r^2 + s^2)} \frac{rs}{\sqrt{r^2 + s^2}} \, ds \, d\phi \\
 &= \frac{Er}{4\pi} \int_0^{2\pi} d\phi \int_0^\infty \frac{s}{(r^2 + s^2)^{\frac{3}{2}}} \, ds \\
 &= \frac{Er}{4\pi} 2\pi \left( \frac{-1}{\sqrt{r^2 + s^2}} \Big|_0^\infty \right) \\
 &= \frac{E}{2}. \tag{8}
 \end{aligned}$$

The result given by equation (8) makes sense. Since it is an infinite virtual plane, one would expect that exactly half of the energy output would eventually reach it.

#### 4.3. Energy density at the receiver using Huygens' principle

With the previous results at hand, we calculate the energy density at a receiver located at a distance  $L = 2r$  from the source, as depicted in figure 8. To do that, we apply Huygens' principle on the surface of the plane.



**Figure 8.** Schematic showing a receiver (magenta triangle) located on the other side of the virtual plane surface, diametrically opposite to the light source. The red lines represent the wavefronts.

As shown in figure 8, each point on the wavefront reaching the plane can be considered a new source of hemispherical waves (waves propagating back to the source do not exist, as previously pointed out). Thus, one evaluates the ‘density of the energy density’ at the receiver ( $\epsilon_R$ ) from the energy density on the plane ( $\epsilon_D$ ) by taking into account the geometrical spreading and the incidence angle once again:

$$\epsilon_R(r, s) = \frac{\epsilon_D(r, s)}{2\pi(r^2 + s^2)} \frac{r}{\sqrt{r^2 + s^2}} = \frac{E}{8\pi^2(r^2 + s^2)^2} \frac{r^2}{r^2 + s^2}. \quad (9)$$

The factor in the denominator of equation (9) is  $2\pi$  instead of  $4\pi$  because we are assuming hemispherical wavefronts. To compute the energy density at the receiver, we must integrate the contributions of the entire plane:

$$\begin{aligned} \epsilon_R(r) &= \int_S dE = \int_S \epsilon_R(r, s) dS = \int_0^{2\pi} \int_0^\infty \epsilon_R(r, s) s ds d\phi \\ &= \int_0^{2\pi} \int_0^\infty \frac{E}{8\pi^2(r^2 + s^2)^2} \frac{r^2 s}{r^2 + s^2} ds d\phi \\ &= \frac{Er^2}{8\pi^2} \int_0^{2\pi} d\phi \int_0^\infty \frac{s}{(r^2 + s^2)^3} ds \\ &= \frac{Er^2}{8\pi^2} 2\pi \left[ \frac{-1}{4(r^2 + s^2)^2} \Big|_0^\infty \right] \\ &= \frac{E}{4\pi(2r)^2}. \end{aligned} \quad (10)$$

Equation (10) retrieves the result from equation (3), in which the energy density at a receiver located at a distance  $L = 2r$  from the source is just the energy output of the source divided by the surface area of the sphere of radius  $L$ . We also evaluate the energy density as a function of

the integration radius  $s$  using a dummy variable of integration  $\xi$ :

$$\begin{aligned}\varepsilon_R(r, s) &= \int_0^{2\pi} \int_0^s \varepsilon_R(r, \xi) d\xi \xi d\phi = \int_0^{2\pi} \int_0^s \frac{E}{8\pi^2(r^2 + \xi^2)^2} \frac{r^2 \xi}{r^2 + \xi^2} d\xi d\phi \\ &= \frac{Er^2}{8\pi^2} 2\pi \left[ \frac{-1}{4(r^2 + \xi^2)^2} \right]_0^s \\ &= \frac{E}{4\pi(2r)^2} \left[ 1 - \frac{r^4}{(r^2 + s^2)^2} \right].\end{aligned}\quad (11)$$

#### 4.4. Simplified model of the sensitivity kernels

To derive a simplified model of the sensitivity kernels, we use the same geometry of the previous example (figures 7 and 8). We also assume a monochromatic source, for simplicity. In this case, the signal  $F(r, t)$  recorded at the receiver can be expressed as:

$$F(r, t) = A_R^*(r) \sin[\omega t + \Phi(r)], \quad (12)$$

where  $A_R^*(r)$  is the amplitude of the wave as a function of  $r$ ;  $\omega = 2\pi f$  is the angular frequency (which in turn is a function of the frequency  $f$ ); and  $\Phi(r)$  is the phase of the wave as a function of  $r$ .

The choice of sine, cosine, or even of the real part of complex exponential in equation (12) is indifferent since they only differ by a phase factor of  $\frac{\pi}{2}$ . Furthermore, given that any pulse can be decomposed into a sum of trigonometric functions using the Fourier transform, it is easy to see how that formulation could be generalized for a multi-frequency light source.

Our goal is to find expressions for  $A_R^*(r)$  and  $\Phi(r)$  as those two functions are crucial to find the approximate amplitude and travel-time kernels we are searching for. By using Huygens' principle again,  $F(r, t)$  can be thought of as the result of the superposition of the rays traveling through each ring of light in figure 7. Therefore:

$$F(r, t) = \int_0^\infty dF = \int_0^\infty A(r, s) \sin[\omega t + \phi(r, s)] ds, \quad (13)$$

where  $dF = A(r, s) \sin[\omega t + \phi(r, s)] ds$  is the contribution of each ring to the resultant signal  $F(r, t)$ ;  $A(r, s)$  is the amplitude of the combined waves coming from each ring as a function of  $r$  and  $s$ ; and  $\phi(r, s)$  is the phase of these same waves, also a function of  $r$  and  $s$ .

The expression for  $\phi(r, s)$  is very straightforward. The phase of each path is directly related to its travel time  $t(r, s)$  by the opposite of the angular frequency:

$$\phi(r, s) = -\omega t(r, s) = -\omega \frac{2d}{c} = -\omega \frac{2\sqrt{r^2 + s^2}}{c}, \quad (14)$$

where  $2d$  is the distance traveled from the source to the receiver as a function of  $r$  and  $s$  (figure 8) and  $c$  is the velocity of propagation. By substituting equation (14) into equation (13) and using the fact that  $\omega = 2\pi f$ , we find:

$$F(r, t) = \int_0^\infty A(r, s) \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right] ds. \quad (15)$$

Finding  $A(r, s)$  requires more effort. In section 4.3, we derived  $\varepsilon_R(r, s)$ , which is the energy density per unit time at the receiver as a function of  $r$  and  $s$ . Finding  $\varepsilon_R(r, s)$  was much easier

than finding the amplitude because we did not have to worry about constructive and destructive interference of the waves. These effects were implicitly taken into account when we combined Huygens' principle and conservation of energy with the effects of geometrical spreading and refocusing at the receiver.

Now, we want to derive an expression for  $A(r, s)$  from  $\varepsilon_R(r, s)$ . However, in equation (15), we explicitly included a sine function whose role is to account for the interference effects. That is crucial to compute the Fresnel zones in our model. However, it also affects  $A_R^*(r)$ , which is the final amplitude measured at the receiver. Nevertheless, since  $\varepsilon_R(r, s)$  already includes the interference implicitly, if we calculate  $A(r, s)$  from it and plug the result directly into equation (15), we would be taking the interference into account twice, finding a wrong expression for the amplitude, thus violating the conservation of energy. Therefore,  $A(r, s)$  must include a renormalization factor to fix this:

$$A(r, s) = R_N A^\circ(r, s), \quad (16)$$

where  $R_N$  is the renormalization factor and  $A^\circ(r, s)$  is the non-renormalized amplitude of the waves coming from each ring.

The energy carried by a wave is directly proportional to the square of its amplitude. Therefore:

$$A_R(r, s) \propto \sqrt{\varepsilon_R(r, s)}, \quad (17)$$

where  $A_R(r, s)$  is the amplitude recorded at the receiver.

For mechanical waves, the constant of proportionality that relates these two quantities also depends on the square of the angular frequency and other physical parameters that vary according to the kind of wave. The energy of a mechanical shear wave (e.g. an  $S$  wave), for example, also depends on the shear modulus (potential energy) and the density of the medium (kinetic energy). Since we are only interested in the amplitude, let us simply assume that these other parameters are implicitly contained in the value of  $E$ . Thus:

$$A_R(r, s) = \sqrt{\frac{E}{4\pi(2r)^2} \left[ 1 - \frac{r^4}{(r^2 + s^2)^2} \right]} = \frac{A_0}{4\pi(2r)} \sqrt{1 - \frac{r^4}{(r^2 + s^2)^2}}, \quad (18)$$

where  $A_0 = \sqrt{4\pi E}$ .

It is easy to see that when  $s \rightarrow \infty$ , equation (18) reduces to the far-field approximation for the amplitude (Aki and Richards 2002). In the far-field, the amplitude at the receiver decays linearly with the distance from the source  $L = 2r$ :

$$A_R(r) = \frac{A_0}{4\pi(2r)}. \quad (19)$$

By setting  $A_0 = 4\pi$  in equation (18) (for simplicity), we find:

$$A_R(r, s) = \frac{1}{2r} \sqrt{1 - \frac{r^4}{(r^2 + s^2)^2}}, \quad (20)$$

where  $A_R(r, s)$  is the amplitude recorded at the receiver as a function of  $r$  and  $s$ .

It is important to highlight that  $A_R(r)$  is not the function  $A_R^*(r)$  we are looking for because it does not account for the effects of interference explicitly. If we tried to derive the expression for the amplitude kernel directly from equation (20), we would not see the Fresnel zones, for example. However, since  $A_R(r, s)$  correctly describes the effects of geometrical spreading

and refocusing of the amplitudes at the receiver, it can be used to find  $A^\circ(r, s)$ .  $A_R(r, s)$  is the amplitude resulting from integrating the contributions from all the rings of light (figure 7). Hence,  $A^\circ(r, s)$ , which is the contribution from each ring, is the partial derivative of  $A_R(r, s)$  with respect to  $s$ :

$$A^\circ(r, s) = \frac{\partial A_R}{\partial s} = \frac{r^3}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}}. \tag{21}$$

By plugging in equation (21) into equation (16) and substituting into equation (15) results in:

$$\begin{aligned} F(r, t) &= \int_0^\infty R_N \frac{r^3}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right] ds \\ &= \int_0^\infty \frac{R_N r^3 \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} ds. \end{aligned} \tag{22}$$

To extract the amplitude  $A_R^*(r)$  from  $F(r, t)$ , we begin by squaring both the left and the right-hand side of equation (12) and integrating in time over the period  $T = \frac{1}{f}$  to eliminate the dependence on  $t$ :

$$\begin{aligned} \int_0^T F(r, t)^2 dt &= \int_0^T \{A_R^*(r) \sin[\omega t + \Phi(r)]\}^2 dt \\ &= A_R^*(r)^2 \int_0^T \sin^2 \left[ \frac{2\pi t}{T} + \Phi(r) \right] dt \\ &= A_R^*(r)^2 \left\{ \frac{t}{2} - \frac{\sin \left[ \frac{4\pi t}{T} + 2\Phi(r) \right]}{8\pi} \right\} \Bigg|_0^T \\ &= A_R^*(r)^2 \frac{T}{2}. \end{aligned} \tag{23}$$

By multiplying both sides of equation (23) by  $\frac{2}{T}$  and taking the square root we find:

$$A_R^*(r) = \sqrt{\frac{2}{T} \int_0^T F(r, t)^2 dt}, \tag{24}$$

where  $F(r, t)$  is given by equation (22).

The final step to complete the derivation of  $A_R^*(r)$  is to determine the renormalization factor  $R_N$ . As mentioned, the role of  $R_N$  is to remove the implicit effects of interference in  $A^\circ(r, s)$ . We begin by using equation (20) to determine how large the integration radius  $s$  has to be so that  $A_R(r, s)$  is 99% of the amplitude  $A_R(r, s \rightarrow \infty)$ :

$$\frac{A_R(r, s)}{A_R(r, s \rightarrow \infty)} = 0.99 \Rightarrow \frac{\frac{1}{2r} \sqrt{1 - \frac{r^4}{(r^2 + s^2)^2}}}{\frac{1}{2r}} = \sqrt{1 - \frac{r^4}{(r^2 + s^2)^2}} = 0.99. \tag{25}$$

By squaring both sides of equation (25) and rearranging, we find a biquadratic equation in  $s$ :

$$\frac{r^4}{(r^2 + s^2)^2} = 1 - 0.99^2 \Rightarrow s^4 + 2r^2s^2 - \left(\frac{0.99^2}{1 - 0.99^2}\right)r^4 = 0. \quad (26)$$

Equation (26) has only one real-positive solution given by:

$$s_{99} = r\sqrt{\frac{100}{\sqrt{199}} - 1}, \quad r \geq 0, \quad (27)$$

where  $s_{99}$  is the integration radius as a function of  $r$  required to achieve 99% of the amplitude  $A_R(r, s \rightarrow \infty)$ . For instance, when  $r = 0.5$ ,  $s_{99} \approx 1.234$ . The choice of 99% was somewhat arbitrary but equation (27) can be generalized to any desired percentage as:

$$s_p = r\sqrt{\frac{1}{\sqrt{1 - \left(\frac{p}{100}\right)^2}} - 1}, \quad r \geq 0, \quad (28)$$

where  $s_p$  is the integration radius as a function of  $r$  required to achieve a percentage  $0 \leq p \leq 100$  of the amplitude. The most important takeaway from equation (28) is that the integration radius required to achieve any percentage is directly proportional to  $r$ .

Section 3.2 briefly explains Fermat's principle as a consequence of constructive and destructive interference of waves taking all possible paths from the source to the receiver. The shorter the wavelength  $\lambda$ , the smaller the deviation from the stationary path required to enter the destructive interference zone (second Fresnel zone) and, consequently, the narrower the kernel. Hence, the wavelength defines how much each path contributes to the final amplitude  $A_R^*(r)$ . On the other hand, the conservation of energy requires that  $A_R^*(r)$  should not change with  $\lambda$ , only with  $r$ . However, without renormalization, the width of the first Fresnel zone would affect  $A_R^*(r)$  in equation (22) by canceling out the contribution of the waves that take a longer path. Any wave traveling outside the first Fresnel zone has little to no impact on  $A_R^*(r)$ . Therefore, to keep the amplitude independent from the wavelength, any increase in  $s_p$  must be accompanied by an equal increase in the width of the first Fresnel zone. However, because equation (28) tells us that  $s_p \propto r$ , if  $\lambda$  changes, the amplitude scales by:

$$S_C = \frac{W_F}{r}, \quad (29)$$

where  $S_C$  is the scaling factor and  $W_F$  is the width of the first Fresnel zone.

Looking at figure 8 and following Snieder (1999), the detour  $D_T$  of a scattered wave relatively to the path taken by the direct wave is:

$$\begin{aligned} D_T &= 2\sqrt{r^2 + s^2} - 2r = 2r\sqrt{1 + \left(\frac{s}{r}\right)^2} - 2r \\ &\approx 2r\left[1 + \frac{1}{2}\left(\frac{s}{r}\right)^2\right] - 2r, \quad s \ll r \\ &\approx \frac{s^2}{r}. \end{aligned} \quad (30)$$

The edge of the first Fresnel zone occurs when  $D_T = \frac{\lambda}{4}$ . Therefore:

$$D_T = \frac{\lambda}{4} \Rightarrow \frac{s^2}{r} = \frac{\lambda}{4} \Rightarrow s = W_F = \frac{\sqrt{\lambda r}}{2} \Rightarrow S_C = \frac{1}{2} \sqrt{\frac{\lambda}{r}}. \quad (31)$$

To prevent scaling,  $R_N$  has to be the reciprocal of  $S_C$ . Thus:

$$R_N = \frac{1}{S_C} = 2\sqrt{\frac{r}{\lambda}}. \quad (32)$$

By substituting equation (32) into equation (22), we find:

$$F(r, t) = \frac{2}{\sqrt{\lambda}} \int_0^\infty \frac{r^{\frac{7}{2}} \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} ds, \quad (33)$$

and, according to equation (24):

$$A_R^*(r) = 2f \sqrt{\frac{2}{c}} \sqrt{\int_0^T \left[ \int_0^\infty \frac{r^{\frac{7}{2}} \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} ds \right]^2 dt}. \quad (34)$$

**4.4.1. Amplitude sensitivity kernel.** To find an expression for the amplitude sensitivity kernel, we begin by rewriting equation (34) as a function of  $s$ :

$$A_R^*(r, s) = 2f \sqrt{\frac{2}{c}} \sqrt{\int_0^T \left[ \int_0^s \frac{r^{\frac{7}{2}} \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right) \right]}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}} d\xi \right]^2 dt}. \quad (35)$$

We evaluate the integral above numerically, fixing  $c = 1$  for simplicity. Figure 9 shows the effect of the wavelength in the approximation by varying the values of  $f$  and figure 10 demonstrates that this model is an excellent approximation for the far-field whenever  $2r \geq \lambda$ .

We are interested in the partial derivative of equation (35) with respect to  $s$ , which approximates the amplitude kernel  $K_A(r, s)$  using our model. Hence:

$$K_A(r, s) = \frac{\partial A_R^*}{\partial s}(r, s) = 2f \sqrt{\frac{2}{c}} \frac{n(r, s)}{m(r, s)}, \quad (36)$$

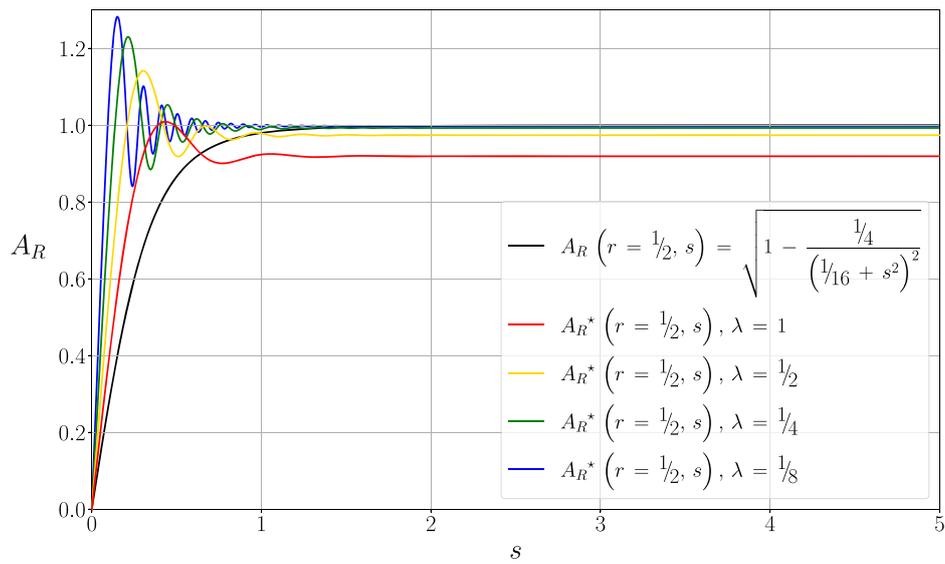


Figure 9. Amplitude as a function of  $s$  for different wavelengths. We set  $r = 1/2$ .

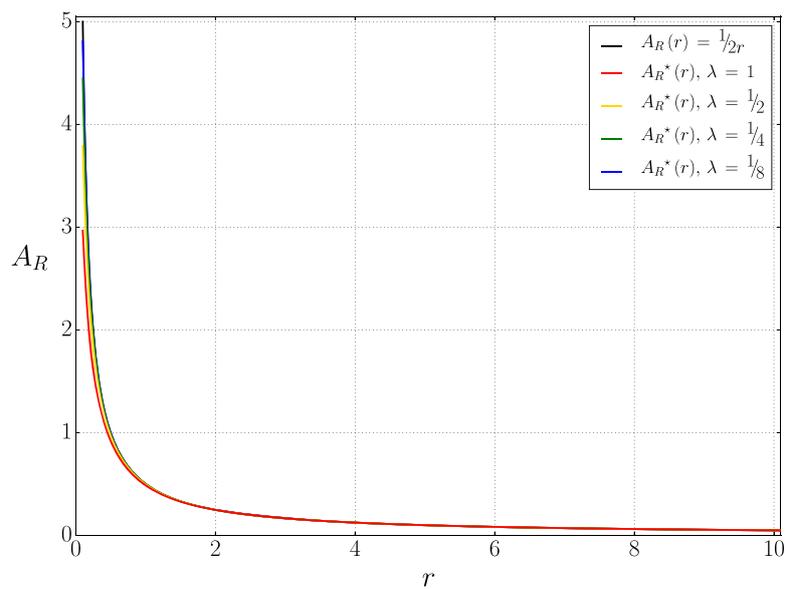


Figure 10. Amplitude decay with  $r$  for different values of  $\lambda$ . The shorter the wavelength, the more the decay rate approaches the linear decay expected in the far-field regime  $A_R(r)$ .

where:

$$n(r, s) = \int_0^T \underbrace{r^{\frac{7}{2}} \sin \left[ \frac{2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right)}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} \right]}_{\alpha(r, s, t)} \times \left\{ \int_0^s \underbrace{\frac{\sin \left[ \frac{2\pi f \left( t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right)}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}} \right]}{\beta(r, s, t)} d\xi \right\} dt$$

and

$$m(r, s) = \sqrt{\int_0^T \left\{ \int_0^s \frac{\sin \left[ \frac{2\pi f \left( t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right)}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}} \right]}{d\xi} d\xi \right\}^2 dt} = \frac{1}{2f} \sqrt{\frac{c}{2}} r^{-\frac{7}{2}} A_R^*(r, s). \tag{37}$$

Equation (36) is not defined for  $s = 0$ . A reasonable solution is replacing  $K_A(r, 0)$  by the limit of  $K_A(r, s)$  when  $s$  approaches zero (see appendix A for the complete derivation):

$$K_A(r, s)|_{s=0} = \lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) = \lim_{s \rightarrow 0} E(r, s) = \sqrt{\frac{2}{\lambda r^3}}, \tag{38}$$

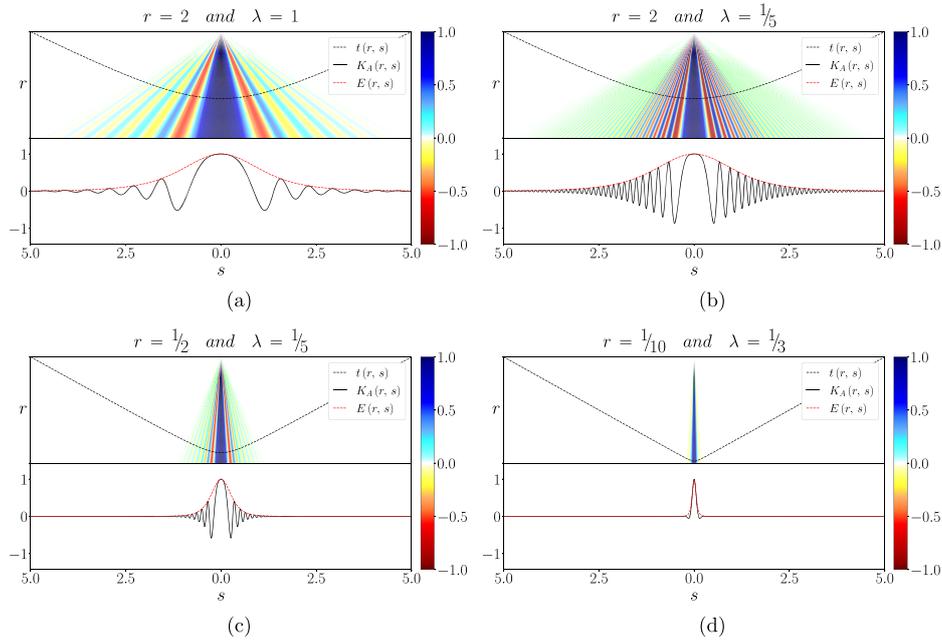
where:

$$E(r, s) = \frac{2r^{\frac{7}{2}}}{(r^2 + s^2)^2 \sqrt{\lambda(2r^2 + s^2)}}. \tag{39}$$

Therefore:

$$K_A(r, s) = \begin{cases} \sqrt{\frac{2}{\lambda r^3}} & \text{if } s = 0, \\ 2f \sqrt{\frac{2}{c}} \frac{n(r, s)}{m(r, s)} & \text{otherwise.} \end{cases} \tag{40}$$

The fact that all paths next to that of the least time have almost identical travel times implies that, for small values of  $s$ , the waves arrive at the receiver with nearly the same phase and constructively interfere, rapidly increasing the amplitude of the detected signal. As  $s$  increases, the waves begin to cancel each other out, resulting in the negative values of  $K_A(r, s)$  in figure 11. From then on, the kernel alternates between regions of constructive and destructive interference, known as Fresnel zones. However, the main contribution to the amplitude at the receiver, by far, comes from the first Fresnel zone, located around the ray path. That is the reason why the maximum amplitude sensitivity is at the ray path.



**Figure 11.** Upper half of the amplitude kernels for different values of  $r$  and  $\lambda$ . The bottom half is just the mirrored version of the upper one. The black dashed line shows the travel-time curve predicted using ray theory. The red dashed line shows function  $E(r, s)$ , which is the envelope of the amplitude sensitivity kernel  $K_A(r, s)$ . The color palette describes the normalized values of  $K_A(r, s)$ , highlighting the Fresnel zones. The kernels' width increase with  $\sqrt{\lambda L}$ .

**4.4.2. Travel-time sensitivity kernel.** So far, we have been studying the effect of the ray paths on the amplitude of the wave detected at the receiver. Now we continue to evaluate how much each path contributes to the travel time of the resultant wave  $F(r, t)$ . The phase is related to the travel time by the opposite of the angular velocity  $-\omega = -2\pi f$ . Equation (14) gives the travel time and phase of each ray path. However, we want the travel time of the detected wave, resultant from the interference of all possible paths. That is directly related to the phase factor  $\Phi(r)$  in equation (12). One can extract  $\Phi(r)$  from  $F(r, t)$  by fixing the value of  $t = t_i$ , normalizing  $F(r, t_i)$  using function  $A_R^*(r)$ , and taking the arcsine of the result. For  $t_i$ , it is convenient to use:

$$t_i = \left( \frac{2r}{\lambda} \bmod 1 \right) T. \quad (41)$$

Whenever the distance between the source and the receiver is a multiple of the wavelength,  $t_i$  is zero. Otherwise,  $t_i$  is a fraction of the period  $T$  corresponding to the remainder of the integer division of the distance  $2r$  by the wavelength  $\lambda$ . That guarantees that we are always evaluating  $F(r, t)$  at the beginning of the cycle, regardless of the distance from the receiver. On the other hand, after the normalization by  $A_R^*(r)$ ,  $F(r, t_i)$  is constrained on the interval  $[-1, 1]$ , which is the domain of the arcsine function. Therefore:

$$\Phi(r) = \arcsin \left[ \frac{F(r, t_i)}{A_R^*(r)} \right]. \quad (42)$$

Now, we convert the  $\Phi(r)$  to travel time by dividing both sides of equation (42) by  $-\omega$  and adding  $\frac{2r}{c}$ , which is the minimum travel time from the source to the receiver, to the right-hand side. Putting everything together, we find the expression for the travel time  $T(r)$ :

$$T(r) = \frac{2r}{c} - \frac{1}{2\pi f} \arcsin \left[ \frac{F(r, t_i)}{A_R^*(r)} \right]. \tag{43}$$

Analogously to equation (35), we also define  $T$  as a function of  $s$ :

$$T(r, s) = \frac{2r}{c} - \frac{1}{2\pi f} \arcsin \left[ \frac{F(r, s, t_i)}{A_R^*(r, s)} \right], \tag{44}$$

where:

$$F(r, s, t) = \frac{2}{\sqrt{\lambda}} \int_0^s \frac{r^{\frac{1}{2}} \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right) \right]}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}} d\xi. \tag{45}$$

Equation (44) is not defined at  $s = 0$  as the factor  $\frac{F(r, s, t_i)}{A_R^*(r, s)}$  results in the indeterminate form  $0/0$ . Again, we replace the value of  $T(r, 0)$  by adopting the limit of  $T(r, s)$  for  $s \rightarrow 0$  (see the complete derivation in appendix B). Therefore:

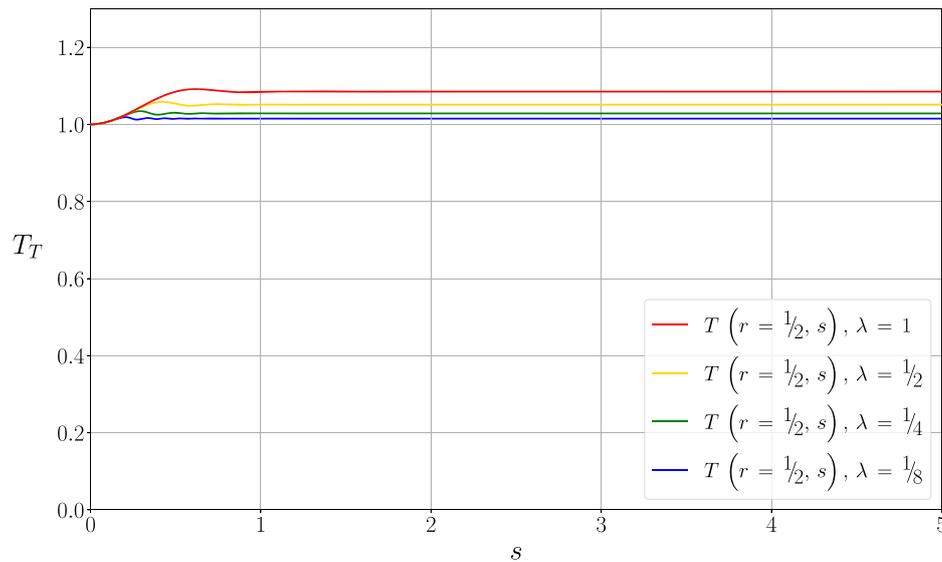
$$T(r, s) = \begin{cases} \frac{2r}{c} & \text{if } s = 0, \\ \frac{2r}{c} - \frac{1}{2\pi f} \arcsin \left[ \frac{F(r, s, t_i)}{A_R^*(r, s)} \right] & \text{otherwise.} \end{cases} \tag{46}$$

The travel-time kernel  $K_T(r, s)$  is given by the partial derivative of (46) with respect to  $s$ :

$$\begin{aligned} \frac{\partial T}{\partial s}(r, s) &= -\frac{1}{2\pi f} \frac{1}{\sqrt{1 - \left[ \frac{F(r, s, t_i)}{A_R^*(r, s)} \right]^2}} \frac{\partial}{\partial s} \left[ \frac{F(r, s, t_i)}{A_R^*(r, s)} \right] \\ &= \frac{1}{2\pi f} \frac{F(r, s, t_i) \frac{\partial A_R^*}{\partial s}(r, s) - \frac{\partial F}{\partial s}(r, s, t_i) A_R^*(r, s)}{A_R^*(r, s)^2 \sqrt{1 - \left[ \frac{F(r, s, t_i)}{A_R^*(r, s)} \right]^2}}. \end{aligned} \tag{47}$$

Once again, we found an expression that is not defined for  $s = 0$ . However, in figure 12 it is easy to see that as  $s$  approaches zero, the derivative of  $T(r, s)$  also approaches zero. That can be formally proved by computing the limit of equation (47) for  $s \rightarrow 0$ , as we did for the former expressions. Nevertheless, the complete derivation is too extensive and falls outside the scope of this paper. Thus:

$$K_T(r, s) = \begin{cases} 0 & \text{if } s = 0, \\ \frac{1}{2\pi f} \frac{F(r, s, t_i) \frac{\partial A_R^*}{\partial s}(r, s) - \frac{\partial F}{\partial s}(r, s, t_i) A_R^*(r, s)}{A_R^*(r, s)^2 \sqrt{1 - \left[ \frac{F(r, s, t_i)}{A_R^*(r, s)} \right]^2}} & \text{otherwise,} \end{cases} \tag{48}$$



**Figure 12.** Travel time as a function of  $s$  for different wavelengths. We set  $r = 1/2$ .

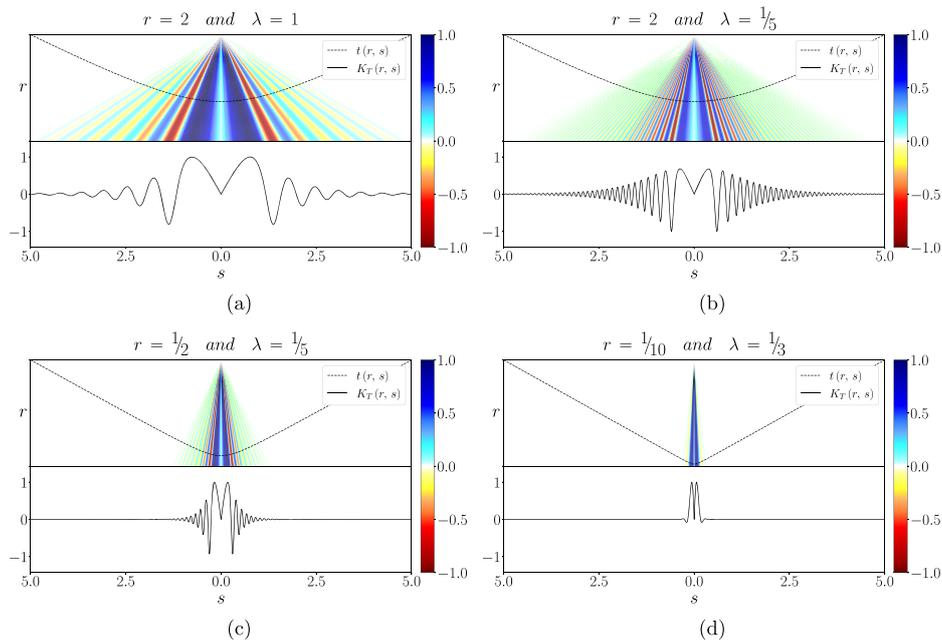
where  $t_i$  is given by equation (41).

In the previous section, the fact that the paths next to the stationary one had nearly identical travel times, making them arrive in phase, explained why the amplitude sensitivity grows rapidly for small values of  $s$ . However, for the very same reason, the travel-time sensitivity in this region is minuscule (figure 13). When all waves arrive aligned, each new wave does not change the phase of the resultant wave, explaining the zero sensitivity at the ray path. As  $s$  increases, the new waves arrive with a progressively larger phase delay. As a result, we enter the destructive interference zone, gradually reducing the amplitude and moving the resultant wave backward. In that region, the travel-time sensitivity reaches its maximum. However, as  $s$  increases even further, the phase difference between the resultant wave and the new arriving waves becomes larger than  $\pi$ . When that happens, the new waves begin to pull the resultant wave back to its original position, originating the negative values of  $K_T(r, s)$ . From then on, the cycle repeats over and over again, originating the higher-order Fresnel zones.

It is important to realize that the larger the phase difference between each new wave and the resultant wave (up to  $\pi$ ), and the smaller the amplitude difference between them, the larger the effect a new wave has on the travel time of the detected signal. The combined effect of these two mechanisms results in the zero sensitivity at the ray path and maximum in the region next (but not very close) to it.

## 5. Colab notebook

A Colab notebook implementing the most important formulas is available at: [https://colab.research.google.com/github/caiciardelli/Sensitivity\\_Kernels/blob/main/Amplitude\\_and\\_travel\\_time\\_sensitivity\\_kernels.ipynb](https://colab.research.google.com/github/caiciardelli/Sensitivity_Kernels/blob/main/Amplitude_and_travel_time_sensitivity_kernels.ipynb).



**Figure 13.** Upper half of the travel-time kernels  $K_T(r, s)$  for different values of  $r$  and  $\lambda$ . The meaning of the symbols and colors is the same as in figure 11, with the difference that there is no equation describing the envelope. In the middle of the kernels (ray path), we observe a region of zero sensitivity.

## 6. Discussions

For simplicity, the model derived in this paper shows the sensitivity with respect to the distance  $s$  from the ray path, not to velocity perturbations, such as the kernels in figure 1. That is why the first Fresnel zones in figures 11 and 13 are positive, unlike the first Fresnel zones in figure 13. Furthermore, the exact computation of the sensitivity kernels would require taking into account **all possible paths** from the source to the receiver, even the craziest ones. In the derivation, only paths that can be decomposed into two straight lines of equal length were considered, which is far more restricted than an accurate computation would require. However, it suffices that this simple approach provided the pivotal conditions required to mimic a more realistic scenario and divert from the straight-line path predicted by ray theory, which would render the model useless. Despite its simplicity, the model correctly predicts the Fresnel zones and the rate with which their width increases with  $\sqrt{\lambda L}$ . The fact that the amplitude decays linearly with the distance came as no surprise since we imposed this during the derivation of the equations. It is also worth noting that the approximation used in this paper could be made progressively more accurate by adding more virtual planes to the formulation. If instead of one, we use two parallel planes, for example, then the rays would have much more degrees of freedom to bend. With infinitely many planes spanning the entire space, we would be considering all possible paths.

## 7. Conclusion

Despite its limitations, the formulation here presented correctly predicts the most relevant features of sensitivity kernels, including the fact that travel-time kernels are hollow. The

step-by-step derivation and relatively easy equations should be understandable by an undergraduate student with a reasonable knowledge of classical physics and calculus. This work provides new insight into the banana-doughnut paradox, helping Earth scientists understand the finite-frequency theory, increasingly used in seismology and seismic exploration.

### Acknowledgments

I thank FAPESP (Grants: 2016/03120-5 and 2018/04918-6) for providing financial support for this research. I also thank three anonymous reviewers for constructive feedback which greatly improved the text and straightened out the exposition.

### Appendix A. Amplitude kernel at the ray path

Here we show the complete derivation to find a representative value for the amplitude kernel when  $s = 0$ , which corresponds to the ray path.

We know that:

$$K_A(r, s)|_{s=0} = \lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) = 2f \sqrt{\frac{2}{c}} \lim_{s \rightarrow 0} \frac{n(r, s)}{m(r, s)} = \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{n(r, s)}{A_R^*(r, s)}. \quad (49)$$

By substituting  $s = 0$  into equation (49), we get an indeterminate form  $0/0$ . Hence, that limit qualifies for the usage of the L'Hôpital's rule:

$$\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) = \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{\frac{\partial n}{\partial s}(r, s)}{\frac{\partial A_R^*}{\partial s}(r, s)} = \frac{8f^2 r^{\frac{7}{2}}}{c} \frac{\lim_{s \rightarrow 0} \frac{\partial n}{\partial s}(r, s)}{\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s)}. \quad (50)$$

By multiplying both members of equation (50) by  $\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s)$ , we find:

$$\left[ \lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) \right]^2 = \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{\partial n}{\partial s}(r, s), \quad (51)$$

where:

$$\begin{aligned} \frac{\partial n}{\partial s}(r, s) &= \frac{\partial}{\partial s} \int_0^T \alpha(r, s, t) \beta(r, s, t) dt \\ &= \int_0^T \left[ \frac{\partial \alpha}{\partial s} \beta(r, s, t) + \alpha(r, s, t) \frac{\partial \beta}{\partial s} \right] dt, \end{aligned} \quad (52)$$

where  $\alpha(r, s, t)$  and  $\beta(r, s, t)$  are given by equation (37).

By the addition rule for limits:

$$\lim_{s \rightarrow 0} \frac{\partial n}{\partial s}(r, s) = \int_0^T \left\{ \lim_{s \rightarrow 0} \left[ \frac{\partial \alpha}{\partial s} \beta(r, s, t) \right] + \lim_{s \rightarrow 0} \left[ \alpha(r, s, t) \frac{\partial \beta}{\partial s} \right] \right\} dt. \quad (53)$$

From equation (37), it is easy to see that  $s \rightarrow 0 \Rightarrow \beta(r, s, t) \rightarrow 0$  and, albeit less obvious,  $s \rightarrow 0 \Rightarrow \frac{\partial \alpha}{\partial s} \rightarrow 0$  too. Therefore, the first limit of equation (53) vanishes.

For the second term, remembering that  $\frac{\partial}{\partial s} \int_0^s f(\xi) d\xi = f(s)$ :

$$\begin{aligned} \alpha(r, s, t) \frac{\partial \beta}{\partial s} &= \int_0^T \frac{r^{\frac{7}{2}} \sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} \\ &\quad \times \left\{ \frac{\sin \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} \right\} dt \\ &= \int_0^T \frac{r^{\frac{7}{2}} \sin^2 \left[ 2\pi f \left( t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^4 (2r^2 + s^2)} dt \\ &= \frac{r^{\frac{7}{2}} \left[ \sin \left( \frac{8\pi f \sqrt{r^2 + s^2} - 4\pi c}{c} \right) - \sin \left( \frac{8\pi f \sqrt{r^2 + s^2}}{c} + 4\pi \right) \right]}{8\pi f (r^2 + s^2)^4 (2r^2 + s^2)} \\ &= \frac{r^{\frac{7}{2}}}{2f (r^2 + s^2)^4 (2r^2 + s^2)}. \end{aligned} \tag{54}$$

Substituting the result of equation (54) into equation (51):

$$\begin{aligned} \left[ \lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) \right]^2 &= \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{r^{\frac{7}{2}}}{2f (r^2 + s^2)^4 (2r^2 + s^2)} \\ &= \frac{4f}{c} \lim_{s \rightarrow 0} \frac{r^7}{(r^2 + s^2)^4 (2r^2 + s^2)}. \end{aligned} \tag{55}$$

Taking the square root of both sides of equation (55):

$$\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) = \lim_{s \rightarrow 0} \frac{2r^{\frac{7}{2}}}{\underbrace{(r^2 + s^2)^2 \sqrt{\lambda(2r^2 + s^2)}}_{E(r, s)}} \implies K_A(r, s) \Big|_{s=0} = \sqrt{\frac{2}{\lambda r^3}}. \tag{56}$$

### Appendix B. Travel-time function at the ray path

In this appendix, we show how to calculate the limit of  $T(r, s)$  as  $s$  approaches zero, used to replace the singularity at  $s = 0$ .

Again, we begin by applying L'Hôpital's rule:

$$\begin{aligned}
 \lim_{s \rightarrow 0} \frac{F(r, s, t_i)}{A_R^*(r, s)} &= \frac{\lim_{s \rightarrow 0} \frac{\partial F}{\partial s}(r, s, t_i)}{\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s)} \\
 &= \frac{\frac{2}{\sqrt{\lambda}} \lim_{s \rightarrow 0} \frac{r^2 \sin \left[ 2\pi f \left( t_i - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}}}{\sqrt{\frac{2}{\lambda r^3}}} \\
 &= r^5 \sqrt{2} \frac{\sin \left[ 2\pi f \left( t_i - \frac{2r}{c} \right) \right]}{r^5 \sqrt{2}} \\
 &= \sin \left\{ 2\pi f \left[ \left( \frac{2r}{\lambda} \bmod 1 \right) T - \frac{2r}{c} \right] \right\} \\
 &= \sin \left\{ 2\pi \left[ \left( \frac{2r}{\lambda} \bmod 1 \right) - \frac{2r}{\lambda} \right] \right\}. \tag{57}
 \end{aligned}$$

Noticing that  $\frac{2r}{\lambda} \bmod 1 = \frac{2r}{\lambda} - \left\lfloor \frac{2r}{\lambda} \right\rfloor$ , where  $\left\lfloor \cdot \right\rfloor$  denotes a floor division, we have:

$$\begin{aligned}
 \lim_{s \rightarrow 0} \frac{F(r, s, t_i)}{A_R^*(r, s)} &= \sin \left[ 2\pi \left( \frac{2r}{\lambda} - \left\lfloor \frac{2r}{\lambda} \right\rfloor - \frac{2r}{\lambda} \right) \right] \\
 &= \sin \left( -2\pi \left\lfloor \frac{2r}{\lambda} \right\rfloor \right) \\
 &= -\sin(2\pi k), \quad k = 0, 1, 2, \dots \infty \Rightarrow \lim_{s \rightarrow 0} \frac{F(r, s, t_i)}{A_R^*(r, s)} = 0. \tag{58}
 \end{aligned}$$

## ORCID iDs

Caio Ciardelli  <https://orcid.org/0000-0001-5316-5886>

## References

- Aki K, Christofferson A and Husebye E S 1977 Determination of the three-dimensional seismic structure of the lithosphere *J. Geophys. Res.* **82** 277–96
- Aki K and Lee W H K 1976 Determination of three-dimensional velocity anomalies under a seismic array using first *P* arrival times from local earthquakes: I. A homogeneous initial model *J. Geophys. Res.* **81** 4381–99
- Aki K and Richards P G 1980 *Quantitative Seismology: Theory and Methods* (San Francisco: Freeman)
- Aki K and Richards P G 2002 *Quantitative Seismology* (Mill Valley, CA: University Science Books)
- Arora K, Cazenave A, Engdahl E R, Kind R, Manglik A, Roy S, Sain K and Uyeda S 2011 *Encyclopedia of Solid Earth Geophysics* (Berlin: Springer)
- Bozdağ E, Trampert J and Tromp J 2011 Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements *Geophys. J. Int.* **185** 845–70

- Brokesčová J 2006 *Asymptotic Ray Method in Seismology: A Tutorial* (Praha: Matfyzpress)
- Cerveny V 2005 *Seismic Ray Theory* (Cambridge: Cambridge University Press)
- Dai D-C and Stojkovic D 2013 Origin of the tail in Green's functions in odd-dimensional space-times *Eur. Phys. J. Plus* **128** 122
- Durand S, Debayle E, Ricard Y, Zanolli C and Lambotte S 2017 Confirmation of a change in the global shear velocity pattern at around 1000 km depth *Geophys. J. Int.* **211** 1628–39
- Dziewonski A M and Anderson D L 1981 Preliminary reference Earth model *Phys. Earth Planet. Inter.* **25** 297–356
- Dziewonski A M, Hager B H and O'Connell R J 1977 Large-scale heterogeneities in the lower mantle *J. Geophys. Res.* **82** 239–55
- Feynman R P 2006 *QED: The Strange Theory of Light and Matter* vol 90 (Princeton, NJ: Princeton University Press)
- French S W and Romanowicz B A 2014 Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography *Geophys. J. Int.* **199** 1303–27
- Grand S P 2002 Mantle shear-wave tomography and the fate of subducted slabs *Phil. Trans. R. Soc. A* **360** 2475–91
- Groenenboom J and Snieder R 1995 Attenuation, dispersion, and anisotropy by multiple scattering of transmitted waves through distributions of scatterers *J. Acoust. Soc. Am.* **98** 3482–92
- Houser C, Masters G, Shearer P and Laske G 2008 Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms *Geophys. J. Int.* **174** 195–212
- Hung S-H, Dahlen F A and Nolet G 2001 Wavefront healing: a banana-doughnut perspective *Geophys. J. Int.* **146** 289–312
- Ishimaru A 1978 *Wave Propagation and Scattering in Random Media* vol 2 (New York: Academic)
- Kennet B L N 1991 IASPEI 1991 seismological tables *Terra Nova* **3** 122
- Kennett B L N, Engdahl E R and Buland R 1995 Constraints on seismic velocities in the Earth from traveltimes *Geophys. J. Int.* **122** 108–24
- Koelemeijer P, Ritsema J, Deuss A and Van Heijst H-J 2016 SP12RTS: a degree-12 model of shear-and compressional-wave velocity for Earth's mantle *Geophys. J. Int.* **204** 1024–39
- Komatitsch D and Tromp J 2002a Spectral-element simulations of global seismic wave propagation: I. Validation *Geophys. J. Int.* **149** 390–412
- Komatitsch D and Tromp J 2002b Spectral-element simulations of global seismic wave propagation: II. Three-dimensional models, oceans, rotation and self-gravitation *Geophys. J. Int.* **150** 303–18
- Lei W *et al* 2020 Global adjoint tomography—model GLAD-M25 *Geophys. J. Int.* **223** 1–21
- Lu C, Grand S P, Lai H and Garnero E J 2019 TX2019slab: a new *P* and *S* tomography model incorporating subducting slabs *J. Geophys. Res. Solid Earth* **124** 11549–67
- Marquering H, Dahlen F A and Nolet G 1999 Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox *Geophys. J. Int.* **137** 805–15
- Marquering H, Nolet G and Dahlen F A 1998 Three-dimensional waveform sensitivity kernels *Geophys. J. Int.* **132** 521–34
- Montelli R, Nolet G, Dahlen F A and Masters G 2006 A catalogue of deep mantle plumes: new results from finite-frequency tomography *Geochem., Geophys., Geosyst.* **7** 11
- Nolet G 2008 *A Breviary of Seismic Tomography: Imaging the Interior of the Earth and Sun* (Cambridge: Cambridge University Press)
- Nolet G, Dahlen F A and Montelli R 2005 Traveltimes and amplitudes of seismic waves: a reassessment *Seismic Earth: Array Analysis of Broadband Seismograms (Geophysical Monograph* vol 157) ed A Levander and G Nolet (Washington, DC: American Geophysical Union) pp 37–47
- Ritsema J, Deuss A, Van Heijst H J and Woodhouse J H 2011 S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements *Geophys. J. Int.* **184** 1223–36
- Ritsema J, van Heijst H J and Woodhouse J H 1999 Complex shear wave velocity structure imaged beneath Africa and Iceland *Science* **286** 1925–8
- Robinson E A and Clark D 2017 *Basic Geophysics* (Tulsa: Society of Exploration Geophysicists)
- Schaeffer A J and Lebedev S 2013 Global shear speed structure of the upper mantle and transition zone *Geophys. J. Int.* **194** 417–49
- Sengupta M K and Toksöz M N 1977 Three dimensional model of seismic velocity variation in the Earth's mantle *Geophys. Res. Lett.* **3** 84–6
- Snieder R 1999 Imaging and averaging in complex media *Diffuse Waves in Complex Media* (Berlin: Springer) pp 405–54

- Snieder R and Lomax A 1996 Wavefield smoothing and the effect of rough velocity perturbations on arrival times and amplitudes *Geophys. J. Int.* **125** 796–812
- Spetzler J and Snieder R 2004 The Fresnel volume and transmitted waves *Geophysics* **69** 653–63
- Tarantola A 1987 *Inverse Problem Theory and Methods for Model Parameter Estimation* (Philadelphia: Society for Industrial and Applied Mathematics)
- Tian Y, Montelli R, Nolet G and Dahlen F A 2007 Computing traveltimes and amplitude sensitivity kernels in finite-frequency tomography *J. Comput. Phys.* **226** 2271–88
- Woodward M J 1992 Wave-equation tomography *Geophysics* **57** 15–26