OPEN

Huygens' Principle geometric derivation and elimination of the wake and backward wave

Forrest L. Anderson

Huygens' Principle (1678) implies that every point on a wave front serves as a source of secondary wavelets, and the new wave front is the tangential surface to all the secondary wavelets. But two problems arise: portions of wavelets that exist outside of the new wave front combine to form a wake. Also there are two tangential surfaces so wave fronts are propagated in both the forward and backward directions. These problems have not previously been resolved by using a geometrical theory with impulsive wavelets that are in harmony with Huygens' geometrical description. Doing so would provide deeper understanding of and greater intuition into wave propagation, in addition to providing a new model for wave propagation analysis. The interpretation, developed here, of Huygens' geometrical construction shows Huygens' Principle to be correct: as for the wake, the Huygens' wavelets disappear when combined except where they contact their common tangent surfaces, the new propagating wave fronts. As for the backward wave, a source propagates both a forward wave and a backward wave when it is stationary, but it propagates only the forward wave front when it is advancing with a speed equal to the propagation speed of the wave fronts.

Huygens' Principle implies that every point on a wave front serves as a source of secondary wavelets. The new wave front is the tangential surface to all the secondary wavelets in the direction of propagation (Fig. 1 presents this as a geometric construction).

An abundance of literature refers to Huygens' Principle^{1–3}. References^{4–32} are a very small sample which show the tremendous variation in the application of Huygens' Principle and some continuing work on it. But there has been insufficient theory developed which addresses the underlying Huygens' basic geometrical construction of how a wave propagates via Huygens' wavelets. "At present mathematical physicists are busy with the rapid developments of modern physics, and older problems are likely to be neglected" (referring to Huygens Principle in 1940)³³.

This is despite the common use of Huygens' construction to introduce wave theory, reflection, refraction, and diffraction. In his construction two problems arise: the portions of the wavelets not contacting the position of the new propagating wave must disappear ('no wake')^{2,3}. Also, the portions of the wavelets propagating in the direction opposite to the new propagating wave front must vanish ('no backward wave'). There has been no clear direct geometrical theory with impulsive wavelets that are in harmony with Huygens' geometrical description which shows how this can occur.

Huygens' Principle's problems of the backward wave and the wake have been approached before by relying on physics (for example, suitable initial values of the velocity potential and condensation³), or by involving doublets, dipoles, obliquity factors, sinusoidal waves, and also by using complex mathematics^{2,3}, "...but the intuitive appeal of Huygens' simple principle is lost..."⁴ along with a clear direct connection to his geometrical construction and the physical insight that derives from it.

Wave propagation is linear so superposition holds: it should be possible to decompose an impulsive propagating wave front into its constituent points, then consider the impulsive wavelets radiating from each of those points at a future time, and combine those wavelets in a simple direct geometric manner to obtain the progressing wave front at that future time.

The following shows that Huygens' geometrical construction is correct as depicted in his figures when the wavelets are interpreted as Dirac Delta distributions (Delta functions)^{34–38}, and when the propagating wave field resulting from the summation of the wavelets is differentiated to yield the wave front, and also when the speed of the source is taken into account. Given this, the backward wave front and wake both vanish.

The following development is geometric and does not involve the wave equation or any particular type of field (except in the Supplementary Information Notes). The symbol φ is used for the propagating wave fields

Unaffiliated, Tucson, USA. email: iimaging@aol.com

Figure 1. Huygens' Principle as a geometrical construction. (a) A replica of Huygens' original Figure 1 (a spherical wave). (b) Huygens' geometrical construction presented as a plane wave.

and f is used for wave fronts derived from φ . The symbol δ is used for the Dirac delta distribution and δ_f is used for finite approximations to it. The symbol θ is used for the Heaviside step function^{36, 37}. All results are in the sense of distributions.

Results

The ensuing analyzes the wave fields and wave fronts from stationary (section “Stationary planar source and elimination of the wake”) and moving (section “Moving planar source and elimination of the backward wave”) impulsive infinite planar sources and also nonexpanding (section “Non expanding spherical source and elimination of the wake”) and expanding (section “Expanding spherical source and elimination of the backward wave”) impulsive spherical sources. This is done because planar and spherical sources are the most fundamental extended wave sources, and the analysis here must work (replicate the original source wave with no wake) for each of them. Also it must produce the correct results (no backward wave) when the sources are moving or expanding at the wave propagation speed.

In addition the planar source analysis supplies necessary results that are required for the spherical source analysis. Also, the planar source results are used to check the spherical source results in the limit as the radius goes to infinity.

Each of the sections “Moving planar source and elimination of the backward wave”, “Non expanding spherical source and elimination of the wake”, “Expanding spherical source and elimination of the backward wave”) depends on, and/or corroborates, results from the preceding sections. In addition the preceding sections are used to corroborate the preceding sections.

The results from section “Stationary planar source and elimination of the wake” are needed to produce the results in section “Moving planar source and elimination of the backward wave”. (This is because in the limit as the pulse width goes to zero the area of the frustum is the same as the area of the annulus) Also in the limit of no source motion section “Moving planar source and elimination of the backward wave” results corroborate section “Stationary planar source and elimination of the wake” results.

The results from section “Non expanding spherical source and elimination of the wake” are needed to produce the results in section “Expanding spherical source and elimination of the backward wave”. (This is because, again, in the limit as the pulse width goes to zero the area of the frustum is the same as the area of the annulus) Also in the limit of no source expansion section “Expanding spherical source and elimination of the backward wave” results corroborate section “Non expanding spherical source and elimination of the wake” results.

The method used for the moving planar source, section “Moving planar source and elimination of the backward wave”, is used in the expanding spherical source, section “Expanding spherical source and elimination of the backward wave”, to implement the effects of motion on the wavelet point sources initial impulses due to the sphere's radius expanding with a speed v : this is possible because, again, in the limit as the pulse width goes to zero the area of the frustum is the same as the area of the annulus.

The notes in the SI are there to corroborate the results in the main text. There are references to each note at the appropriate location in the main text.

An exception, note E geometrically derives the integral term in D'Alemberts formula which is used in the SI. This derivation is originated here.

With this I have shown the linkage between the analysis sections “Stationary planar source and elimination of the wake”, “Moving planar source and elimination of the backward wave”, “Non expanding spherical source and elimination of the wake”, “Expanding spherical source and elimination of the backward wave”, and also the linkage with the notes in the SI.

Each major analysis section derives a major conclusion:

Section I: a stationary planar source has no wake.

Section II: a planar source moving at the speed of propagation also has no backward wave.

Section III: a non expanding spherical source has no wake.

Section IV: an spherical source expanding at the speed of propagation also has no backward wave.

Figure 2. Planar source. (a) An infinite plane with impulsive excitation, $\delta_f(t)$. The excitation duration or temporal pulse width is ε . The parameter z_0 is the distance from the plane to a point in space, P , at which the wave field will be observed as a function of time. The propagation speed is c , where $c > 0$. R_1 and R_2 are the radii of an annulus defined by the intersection of the plane with two spheres of radii ct and $c[t + \varepsilon]$. The point N is the nearest point on the plane to the field point P . Each point on the plane is a point source which radiates an impulsive spherical wavelet. (b) The wave field ϕ that is propagated from the plane. (c) The wave front f in the $Z+$ direction that is derived from the wave field which will arrive at P at $t = z_0/c$.

Note that the moving sources (planar or spherical) moving at the speed of propagation, c , are in fact the propagating wave fronts populated by secondary sources and mentioned in Huygens' principle; and are the reason for the backward wave elimination, as will be shown..

Stationary planar source and elimination of the wake. (The infinite planar source may sometimes be referred to as 'planar source' or 'plane').

This section analyzes the waves propagating away from a motionless infinite planar source. Each point on the plane is a source of a Huygens' wavelet consisting of a spherical Dirac delta function (distribution) expanding as a function of time at a rate equal to c , the propagation speed. The strength (or amplitude) of the wavelet is α which is attenuated as the wavelet expands by spherical spreading.

It is shown that the wave field (summation of the Huygens' wavelets) observed as a function of time at some fixed point is a constant value after the wavelets first reach that point. The partial derivative with respect to time of that wave field gives the wave front which replicates the original wave (the Huygens' wavelet) at $1/2$ amplitude in both the advancing and retreating directions. The derivative also shows there is no wake. See Fig. 2.

It is only necessary to derive the wave field for each point on any one line perpendicular to the source since the planar source is uniform with respect to the x and y coordinates. Only the points in the planar source intersected by the sphere with radius ct centered at P can contribute to the wave field $\phi(z_0, t)$ when deriving the wave field at P at time t (Fig. 2a).

Initially the duration ε of the excitation pulse δ_f will be finite and the pulse will have an height equal to α . Then the area (or strength) under this finite version of $\delta_f(t)$ is $\varepsilon\alpha$ which is required to be 1. Later, where needed, the limit as ε tends toward zero will be taken while still requiring α to take on a value such that $\varepsilon\alpha$ remains equal to 1 (i.e. $\alpha = 1/[\varepsilon c]$) so that $\delta_f \rightarrow \delta$ as $\varepsilon \rightarrow 0$.

The locus of all radiating points on the planar source which can contribute to the wave field $\phi(z_0, t)$ during the time interval $[t, t + \varepsilon]$ and at distance z_0 from the plane is an annular area. This annular area is bounded by two concentric spherical shells of radius ct and $c[t + \varepsilon]$.

The area A_p of the annulus is:

$$A_p = \pi R_2^2 - \pi R_1^2 \quad (1)$$

where

$$R_1^2 = c^2 t^2 - z_0^2 \quad (2)$$

and

$$R_2^2 = c^2 [t + \varepsilon]^2 - z_0^2 = c^2 [t^2 + 2t\varepsilon + \varepsilon^2] - z_0^2 = c^2 t^2 + 2c^2 t\varepsilon + c^2 \varepsilon^2 - z_0^2 \quad (3)$$

So

$$A_p = \pi c^2 [2t\varepsilon + \varepsilon^2] \quad (4)$$

which has the limit

$$A_p = 2\pi c^2 t\varepsilon \text{ as } \varepsilon \rightarrow 0 \quad (5)$$

(and so $\delta_f \rightarrow \delta$).

Spherical spreading attenuation specified to be of the form $1/[4\pi ct]$ attenuates the wavelets and therefore the radiation from the annular area. As a result the attenuated radiation from Ap is not a function of the field point's distance, z_0 , from the planar source (except for a step at $ct = z_0$ as the wave front passes). This will be shown in the following:

Combining the $1/[4\pi ct]$ spherical spreading attenuation and the α height factor with the area Ap gives the (dimensionless) amplitude of the wave field, φ , after the impulsive spherical wavelets have all passed (i.e. for $|z| \leq ct$), yielding:

$$\varphi = Ap \frac{\alpha}{4\pi ct} = \frac{2\pi c^2 t \varepsilon}{4\pi ct} \alpha = \frac{2\pi ct}{4\pi ct} c\varepsilon\alpha = \frac{c\varepsilon\alpha}{2} \quad (6)$$

Because $c\varepsilon\alpha = 1$, the result simplifies to

$$\varphi = \frac{1}{2} \text{ for } |z| \leq ct \quad (7)$$

or using a Heaviside step function, θ ,

$$\varphi(z, t) = \frac{1}{2} \theta(ct - |z|) \quad (8)$$

which as a spatial function of z is a rectangular pulse expanding in time in both advancing (toward P) and retreating (away from P) directions (Fig. 2b).

In the advancing direction the wave field observed at P as a function of time is

$$\varphi(z_0, t) = \frac{1}{2} \theta(ct - z_0) \quad (9)$$

which has a transition at $t = z_0/c$ when the wave field has advanced to P .

The derivative of φ yields an advancing planar wave front, f , which observed at P as a function of time is (Fig. 2c):

$$f^+(z_0, t) = \frac{\partial}{\partial t} \frac{1}{2} \theta(ct - z_0) = \frac{c}{2} \delta(ct - z_0) = \frac{c}{2} \delta\left(c\left[t - \frac{z_0}{c}\right]\right) = \frac{c}{2c} \delta\left(t - \frac{z_0}{c}\right) = \frac{1}{2} \delta\left(t - \frac{z_0}{c}\right) \quad (10)$$

Or for z in general

$$f^+(z, t) = \frac{\partial}{\partial t} \frac{1}{2} \theta\left(t - \frac{z}{c}\right) = \frac{1}{2} \delta\left(t - \frac{z}{c}\right) \quad (11)$$

This holds for every line perpendicular to the source. Note that only the single point source at the nearest point N contributes to the wave front at P : the wave field φ is a step function and its derivative is zero except at the step's location, and the step was caused by the wave field originating at N (Fig. 2a). (So there is a one to one correspondence between the points in the wave front and the points in the planar source.) Because the derivative is zero except at the step, there is no wake.

There will also be a retreating plane wave front $f^-(z, t) = \frac{1}{2} \delta(t + z/c)$ propagating in the opposite direction, away from P , which can be derived similarly. This also propagates without a wake. (Here z is negative, and f will not be seen at z_0 , the location of P).

Therefore an infinite planar distribution of point sources radiating impulsive spherical wavelets creates advancing and retreating impulsive plane wave fronts that propagate cleanly without a wake. This shows that the Huygens' wavelets cancel each other when summed together and differentiated except on their common tangential surfaces (Fig. 1a,b).

It might be expected that the previous process can be repeated and the impulsive planar wave fronts made to progress to the next future time $t + \Delta t$, and so on. But because of the $\frac{1}{2}$ factor the waves would be attenuated upon each iteration. When the source is moving at the same speed as the advancing wave front the $\frac{1}{2}$ factor disappears along with the retreating or backward wave front, as will be shown:

Moving planar source and elimination of the backward wave. This section analyzes the changes that occur in the results of section "Stationary planar source and elimination of the wake" when the planar source is moving with speed v toward the observation point. It is shown that the results from section "Stationary planar source and elimination of the wake" can be used if the effects of motion on the Huygens' wavelets are included. When this is done it is seen that a new additional wave is created as a function of the planar source motion, v . This additional wave reduces the amplitude of the backward wave and increases the amplitude of the forward wave. When $v = c$ the backward wave is eliminated and the forward wave amplitude is doubled eliminating the $1/2$ factor. In any case the original wave is replicated. See Fig. 3.

During the time interval ε the planar source moves a distance $v\varepsilon$ in the direction toward P . Then $R2$ is the radius of the base of a right circular cone of slant height $c[t + \varepsilon]$ and $R1$ is the radius of the base of a right circular cone of slant height ct . These two bases define a conical frustum of height $v\varepsilon$ and slant height d . The lateral area of this frustum, $\pi[R2 + R1]d$, is the locus of all radiating points which can contribute to the field at the field point P during the time interval $[t, t + \varepsilon]$. This is analogous to the annular area on the infinite stationary plane previously considered.

Figure 3. The geometry of an infinite planar impulsive source moving toward P with speed v . (a speed away from P would be given by $-v$.) R_1 is the radius of the circle created by the intersection at time t of a sphere of radius ct with the plane. R_2 is the radius of the circle created by the intersection at time $t + \epsilon$ of a sphere of radius $c[t + \epsilon]$ with the plane.

However, in the limit of $\epsilon \rightarrow 0$ the area of the frustum reduces to the area of that annulus as follows:

The annulus area is $\pi[R_2^2 - R_1^2]$ and the frustum area is $\pi[R_2 + R_1]d$. Then referring to Fig. 3, $d^2 = v^2\epsilon^2 + [R_2 - R_1]^2$, and as $\epsilon \rightarrow 0$, $d^2 = [R_2 - R_1]^2$. So $d = R_2 - R_1$.

Therefore as $\epsilon \rightarrow 0$ the frustum area becomes $\pi[R_2 + R_1]d = \pi[R_2 + R_1][R_2 - R_1] = \pi[R_2^2 - R_1^2]$, which is also the area of the annulus.

Consequently the previously developed results and formulas for the infinite stationary plane can be used. However, the effects of the motion of the moving source on its point sources' initial radiated impulses must still be considered:

See Fig. 4a–c. If the planar impulsive source is moving with positive source speed, v , toward P , the motion results in a displacement during the pulse time interval ϵ (an additional initial condition) which causes additional wave front pulses to propagate in both the forward and backward directions, as with the stationary planar source. (Displacement at x is approximately the product of the temporal integration interval and the displacement speed evaluated at x . See Supplementary Notes B and E) This is in addition to the previously derived wave front pulses from the stationary planar impulsive source. The displacement height is the same as the original forward wave front pulse, $\alpha/2$, whether longitudinal or transverse. The displacement spatial extent is $v\epsilon$.

The source motion in the forward direction displaces a positive $v\epsilon\alpha/2$ during the time interval ϵ which propagates forward as an additional wave front pulse. Since $\alpha = 1/[c\epsilon]$ the displacement or strength equals $[1/2][v\epsilon]/[c\epsilon] = v/2c$. Do to linearity this additional wave front pulse is added to the original stationary planar source wave front pulse. Note that as $\epsilon \rightarrow 0$ a pulse of width ϵ and height $1/\epsilon$ moving at speed c tends to $\delta(x - ct)$. Note also that v and c are scalars and do not have to be measured along the same direction—they effect pulse strength by different means.

Then the total resulting wave front in the forward direction is the sum of two wave front pulses which, as $\epsilon \rightarrow 0$, can be combined to yield:

$$f^+(z, t) = \frac{1}{2}\delta\left(t - \frac{z}{c}\right) + \frac{v}{c}\frac{1}{2}\delta\left(t - \frac{z}{c}\right) = \left[1 + \frac{v}{c}\right]\frac{1}{2}\delta\left(t - \frac{z}{c}\right) = \frac{v+c}{2c}\delta\left(t - \frac{z}{c}\right) \quad (12)$$

See Fig. 4d–f. Proceeding similarly with the equalizing additional backward wave where the displacement is $-v\epsilon\alpha/2$ since the propagation direction is opposite to the source motion direction. The source motion displaces an equalizing negative $-v\epsilon\alpha/2$ during the pulse time interval ϵ which then propagates backward as an additional wave front pulse. Since $\alpha = 1/[c\epsilon]$ the displacement or strength equals $[1/2][-v\epsilon]/[c\epsilon] = -v/2c$. As before this wave front pulse is added to the original stationary planar source wave front pulse. Then the total resulting wave front in the backward direction (where z is negative) is the sum of two wave front pulses which, as $\epsilon \rightarrow 0$, can be combined to yield:

$$f^-(z, t) = \frac{1}{2}\delta\left(t + \frac{z}{c}\right) + \frac{-v}{c}\frac{1}{2}\delta\left(t + \frac{z}{c}\right) = \left[1 + \frac{-v}{c}\right]\frac{1}{2}\delta\left(t + \frac{z}{c}\right) = \frac{-v+c}{2c}\delta\left(t + \frac{z}{c}\right) \quad (13)$$

The sign is negative for the additional wave front pulse that is propagating in the direction opposite to the motion of the source (in the backward direction). This relative direction of motion defines the terms 'forward' and 'backward' in the case of the moving planar impulsive source. (Note that the terms are not defined for a stationary planar impulsive source).

So the impulsive forward and backward planar wave fronts are

Figure 4. (a–c) Source motion displacement for the forward wave: (a) the original forward wave front pulse compared with the additional forward wave front pulse due to the source motion. (b) The original forward wave front pulse which is the same as the wave front pulse from the stationary source. (c) The additional forward wave front pulse due to the source motion. (d–f) Source motion displacement for the backward wave: (d) the original backward wave front pulse compared with the additional backward wave front pulse due to the source motion. (e) The original backward wave front pulse which is the same as the wave front pulse from the stationary source. (f) The additional backward wave front pulse due to the source motion. (g–i) The original and additional wave fronts with $v = c$. (g) Forward and backward wave front pulses that would propagate from a stationary source. (h) Additional wave front pulses originating from source motion. (i) Forward wave front pulse created by the sum of the other two wave front pulses (g) and (h).

$$f^+(z, t) = \frac{v+c}{2c} \delta(t - z/c) \text{ and } f^-(z, t) = \frac{-v+c}{2c} \delta(t + z/c) \quad (14)$$

(At P at a positive distance along the z axis, only the forward wave front will be observed.) For the stationary planar source (where $v=0$), we get $f^+(z, t) = \frac{1}{2} \delta(t - z/c)$ for the forward wave front and $f^-(z, t) = \frac{1}{2} \delta(t + z/c)$ for the backward wave front, as was derived before.

See Fig. 4g–i. For source speed $v=c$ the forward wave front is $f^+(z, t) = \delta(t - z/c)$ or as seen at P as a function of time, $f^+(z_0, t) = \delta(t - z_0/c)$. So the forward wave front propagates without changing and becomes the new progressing Huygens' wave front (Note the absence of the $\frac{1}{2}$ factor). For the backward wave front $f^-(z, t) = 0$. So there is no backward wave front (See SI Notes C, D).

Non expanding spherical source and elimination of the wake. This section analyzes the waves propagating away from a non expanding spherical source. Each point on the sphere is a source of a Huygens' wavelet consisting of a spherical Dirac delta function (distribution) expanding as a function of time at a rate equal to c , the propagation speed. The strength (or amplitude) of the wavelet is α which is attenuated as the wavelet expands by spherical spreading.

It is shown that the wave field (summation of the wavelets) observed as a function of time at some fixed point is a constant value after the wavelets first reach that point and remain constant for a period of time that is proportional to the sphere's diameter. The partial derivative with respect to time of that wave field gives the wave front which replicates the original wave (the Huygens' wavelet) at $1/2$ amplitude in both the expanding and converging directions. That derivative also shows there is no wake. See Fig. 5.

Since the spherical source is symmetrical, it is only necessary to derive the wave field for each point on any one radial line emanating from its center. The wave field will be the same on any other radial line. As before, initially it is assumed that the temporal duration ε of the excitation pulse δ_f is finite and that it has a height equal to α and an area equal to $\varepsilon\alpha$, where $\alpha = 1/[\varepsilon c]$ to maintain an area equal to 1 as $\varepsilon \rightarrow 0$ (and $\delta_f \rightarrow \delta$). The surface area of a spherical cap is $2\pi R_0 h(t)$ (Fig. 5a), and the surface area of the spherical zone is $A_s = 2\pi R_0 [h(t+\varepsilon) - h(t)]$ (Fig. 5b); Note surprisingly that the surface area of the zone depends only on its height, $h(t+\varepsilon) - h(t)$, not on its vertical position within the sphere.

To find A_s the heights $h(t)$ and $h(t+\varepsilon)$ must be found. Two right triangles can be formed (Fig. 5a):

$$T1: c^2 t^2 = R_1^2 + [h(t) + z_0]^2 \text{ and} \quad (15)$$

Figure 5. Non expanding spherical shell impulsive source. **(a)** A spherical cap on an impulsive spherical shell source. (The spherical shell may be referred to as a 'sphere' hereafter.) The radius is R_0 , and the sphere is centered at the coordinate origin and has impulsive excitation $\delta_f(t)$. Each point on the sphere's surface represents a point source which radiates an impulsive spherical wavelet. The parameter z_0 is the radial distance from the sphere's surface to the point in space, P , at which the resultant wave field ϕ will be observed as a function of time. The propagation speed is c . The distance ct from P to a point on the sphere defines a spherical cap of height $h(t)$ with radius R_1 . The sphere is transparent to radiation so that the whole sphere will contribute to the wave field at any one point in space. **(b)** The locus of all contributing radiating points is a spherical zone. These are the only points on the spherical source of radius R_0 which can contribute to the field at P during the time interval $[t, t + \epsilon]$ and at distance z_0 from the sphere. The zone on the spherical source is bounded by its intersection with two concentric spheres of radius ct and radius $c[t - \epsilon]$ centered at the field point P . The intersections defines spherical caps of heights $h(t)$ and $h(t + \epsilon)$, also radii R_1 and R_2 .

$$T2: R_0^2 = R_1^2 + [R_0 - h(t)]^2 \quad (16)$$

Solving for $h(t)$ gets

$$T1 - T2 : c^2t^2 - R_0^2 = [h(t) + z_0]^2 - [R_0 - h(t)]^2 \quad (17)$$

$$c^2t^2 - R_0^2 = h^2(t) + 2h(t)z_0 + z_0^2 - R_0^2 + 2R_0h(t) - h^2(t) \quad (18)$$

$$c^2t^2 - R_0^2 = 2h(t)z_0 + z_0^2 - R_0^2 + 2R_0h(t) \quad (19)$$

$$c^2t^2 = 2h(t)z_0 + z_0^2 + 2R_0h(t) \quad (20)$$

$$2h(t)[R_0 + z_0] = c^2t^2 - z_0^2 \quad (21)$$

$$h(t) = \frac{c^2t^2 - z_0^2}{2[R_0 + z_0]} \quad (22)$$

and also (by substituting $t + \epsilon$ for t)

$$h(t + \epsilon) = \frac{c^2[t + \epsilon]^2 - z_0^2}{2[R_0 + z_0]} \quad (23)$$

where both equations are for time t where $z_0/c \leq t \leq [R_0 + z_0]/c$.

Then the zone's surface area is

$$\begin{aligned}
 As &= 2\pi R_0[h(t + \varepsilon) - h(t)] = 2\pi R_0 \frac{c^2[t + \varepsilon]^2 - z_0^2 - c^2t^2 + z_0^2}{2[R_0 + z_0]} \\
 &= 2\pi R_0 \frac{c^2t^2 + c^22t\varepsilon + c^2\varepsilon^2 - z_0^2 - c^2t^2 + z_0^2}{2[R_0 + z_0]} = \pi R_0 \frac{c^22t\varepsilon + c^2\varepsilon^2}{R_0 + z_0}
 \end{aligned} \quad (24)$$

and as $\varepsilon \rightarrow 0$ (so $\delta_f \rightarrow \delta$)

$$As = \frac{2\pi R_0 c^2 t \varepsilon}{R_0 + z_0} = \frac{R_0}{R_0 + z_0} 2\pi c^2 t \varepsilon \quad (25)$$

which is the same as the area of the stationary infinite plane annulus Ap but attenuated by the spherical spreading factor $R_0/[R_0 + z_0]$ (the radial lines represent infinitesimal solid angles). This area is also the same as the area of the plane's annulus Ap when $R_0 \rightarrow \infty$ or when $z_0 \rightarrow 0$.

As before spherical spreading of the form $1/[4\pi ct]$ attenuates the wavelets and therefore the radiation from the spherical zone. The dimensionless amplitude of φ observed at P for $z_0/c \leq t \leq [2R_0 + z_0]/c$ is produced by combining the area As with the attenuation $1/[4\pi ct]$ and the pulse height α :

$$|\varphi| = \frac{\alpha As}{4\pi ct} = \frac{\alpha 2\pi R_0 t \varepsilon c^2}{4\pi ct[R_0 + z_0]} = \frac{R_0 c \varepsilon \alpha}{2[R_0 + z_0]} \quad (26)$$

Because $c\varepsilon\alpha = 1$

$$|\varphi| = \frac{R_0}{2[R_0 + z_0]} \quad (27)$$

which holds for $z_0 \leq ct \leq [2R_0 + z_0]$. (See Supplementary Note A for a different derivation) For fixed z_0 , $|\varphi|$ observed at P is a rectangular pulse in time having temporal width $2R_0/c$. Note that as $z_0 \rightarrow 0$ or $R_0 \rightarrow \infty$, $|\varphi| \rightarrow 1/2$, which is the same as for the planar source.

The wave field φ observed at P as a function of time is produced by using Heaviside step functions to implement the time dependence, $z_0 \leq ct \leq [2R_0 + z_0]$:

$$\varphi(z_0, t) = |\varphi|[\theta(ct - z_0) - \theta(ct - [z_0 + 2R_0])] = \frac{R_0}{2[R_0 + z_0]} [\theta(ct - z_0) - \theta(ct - [z_0 + 2R_0])] \quad (28)$$

The derivative of φ yields the wave front observed at P as a function of time (Fig. 6b):

$$\begin{aligned}
 f(z_0, t) &= \frac{\partial \varphi(z_0, t)}{\partial t} = \frac{cR_0}{2[R_0 + z_0]} [\delta(ct - z_0) - \delta(ct - z_0 - 2R_0)] \\
 &= \frac{cR_0}{2[R_0 + z_0]} \left[\delta\left(c\left(t - \frac{z_0}{c}\right)\right) - \delta\left(c\left(t - \frac{z_0 + 2R_0}{c}\right)\right) \right] \\
 &= \frac{cR_0}{2[R_0 + z_0]} \left[\frac{1}{c} \delta\left(t - \frac{z_0}{c}\right) - \frac{1}{c} \delta\left(t - \frac{z_0 + 2R_0}{c}\right) \right] \\
 &= \frac{R_0}{2[R_0 + z_0]} \left[\delta\left(t - \frac{z_0}{c}\right) - \delta\left(t - \frac{z_0 + 2R_0}{c}\right) \right]
 \end{aligned} \quad (29)$$

For finite z_0 and t in the equation for $f(z_0, t)$, as $R_0 \rightarrow \infty$ the argument of $\delta(ct - z_0 - 2R_0)$ never equals zero, so never 'activates' and $\delta(ct - z_0 - 2R_0)$ may be ignored. Then

$$f(z_0, t) = \frac{R_0}{2[R_0 + z_0]} \delta\left(t - \frac{z_0}{c}\right) = \frac{1}{2} \delta\left(t - \frac{z_0}{c}\right) \text{ as } R_0 \rightarrow \infty \quad (30)$$

which is the same as the equation for the planar source advancing wave front.

Both propagating spherical wave fronts (Fig. 6a) are affected by the spherical spreading term $R_0/[R_0 + z_0]$ which supplies the additional attenuation due to the radius increasing from the initial R_0 to $R_0 + z_0$ (it is implied that spherical spreading attenuation is included in the initial strength of the wavelets originating on the sphere's surface). For example, $[1/R_0] [R_0/[R_0 + z_0]] = 1/[R_0 + z_0]$. In $-\delta(t - [z_0 + 2R_0]/c)$ the inverse spreading due to going from the far surface to the center is exactly cancelled by the spreading resulting from going from the center to the near surface. So the same $R_0/[R_0 + z_0]$ spherical spreading term can ultimately be applied to both the near and far surface waves that are to be observed at P .

The wave field $\theta(ct - z_0 - 2R_0)$ originated as a positive amplitude propagating wave field on the far surface (in the limit as $R_0 \rightarrow \infty$ the wave fields close to the surface of the sphere must be the same as the wave fields from the plane, and the wave fields on both sides of the plane are positive). As the initially converging spherical wave field passes through the center its amplitude changes sign (Gouy phase shift^{39–41}; "It is well known that a spherical converging light wave undergoes a phase change of 180 degrees in passing through its focus...and is, in fact, a general property of any focused wave.", Boyd p. 877)⁴⁰, and see Fig. 6d.

Figure 6. Radiated wave fronts and individual terms in the equations. The factor $R_0/[2[R_0 + z_0]]$ is not included in the figures' annotations. **(a)** A spherical source of radius R_0 with expanding and converging wave fronts. The distance z_0 is from the surface of the spherical source to the field point P at which the wave fronts will be observed. The two points N and F are the nearest and furthest points from P on the near and far surfaces of the spherical source. **(b)** Shows the impulses from the equation for the wave front $f(z_0, t)$: the impulse $\delta(t - z_0/c)$ observed at P , at distance z_0 , is caused by an expanding impulsive spherical wave. This is the forward wave front for the spherical source from its near surface. The impulse $-\delta(t - [z_0 + 2R_0]/c)$ is caused by an impulsive spherical wave front, observed at P , that had initially converged with decreasing radius inward toward the spherical source's center, passed through the center while changing sign, and then expanded with increasing radius toward P . This is the backward wave front for the spherical source from the far surface. **(c)** The wave field corresponding to the forward wave front. **(d)** The wave field corresponding to the backward wave front. **(e)** The combined wave field resulting from the summation of the forward and backward wave fields.

The wave field then begins expanding. When it expands through the surface of the spherical source the now negative wave field cancels the remainder of the wave field that originated on the near surface, $\theta(ct - z_0)$, (Fig. 6c). The combined wave field (Fig. 6e) when observed in time at P is a rectangular pulse of time duration $2R_0/c$.

Notice that only the two points N and F (Fig. 6a), contribute as sources to the wave front $f(z_0, t)$. This is because $\varphi(z_0, t)$ is a rectangular pulse of temporal width $2R_0/c$ for fixed z_0 and its derivative is zero except on its leading and trailing edges (which originated at the two points N and F). Because the derivative is zero there is no wake (except for an impulse corresponding to the trailing edge). Also there is a one to one correspondence between the points on the source and the points on the wave front.

Therefore a spherical distribution of point sources radiating impulsive spherical wavelets creates two impulsive spherical wave fronts (Fig. 6b). Both propagate cleanly without a wake. Again this shows that the Huygens' wavelets cancel each other when summed together and differentiated, except where they make contact with their tangent surfaces.

The same $1/2$ factor occurs in f as occurred in the case of the stationary infinite planar source. Note also there are forward (expanding) and backward (converging) wave fronts. When the sphere's radius expands with a speed equal to the propagation speed, the $1/2$ factor and the backward wave both disappear, as will be shown.

Expanding spherical source and elimination of the backward wave. This section analyzes the changes that occur in the results of section "Non expanding spherical source and elimination of the wake" when the spherical source is expanding with radial speed v . It is shown that the results for the non expanding sphere in section "Non expanding spherical source and elimination of the wake" can be used if the effects of motion on the Huygens' wavelets are included.

When this is done it is seen that a new additional wave is created. (A similar additional wave is discussed in section "Moving planar source and elimination of the backward wave".) The amplitude of this additional wave is a function of the radial expansion speed, v . This additional wave reduces the amplitude of the backward wave and increases the amplitude of the forward wave. When $v = c$ the backward wave is eliminated and the forward wave amplitude is doubled. In any case the original wave is replicated and there is no wake. See Fig. 7.

Figure 7. The impulsive spherical shell source (the 'sphere') with initial radius R_0 expanding with a positive relative radial speed v where as with the planar source, v will be positive for forward wave fronts (in the direction of the expansion) and will be preceded by a negative sign for backward wave fronts. $R1$ is the radius of the circle created by the intersection at time t of a sphere centered at P of radius ct with the spherical source. $R2$ is the radius of the circle created by the intersection at time $t + \epsilon$ of a sphere centered at P of radius $c[t + \epsilon]$ with the spherical source.

See Fig. 7. During the time interval ε the spherical source radius expands a distance $v\varepsilon$. Similar to the case of the moving planar source, R_2 is the radius of the base of a right circular cone of slant height $c[t + \varepsilon]$ and R_1 is the radius of the base of a right circular cone of slant height ct . These two bases, R_1 , R_2 , define a conical frustum of height Δh and slant height d . The lateral area of this frustum, $\pi[R_2 + R_1]d$, is the locus of all radiating points which can contribute to the field at the point P during the time interval $[t, t + \varepsilon]$.

The expanding equivalent, As_e , of the nonexpanding spherical zone area, As , must be derived. This is the lateral area of the frustum.

The frustum area is

$$As_v = \pi[R1 + R2]d \quad (31)$$

where the slant height is

$$d^2 = \Delta h^2 + [R2 - R1]^2 \quad (32)$$

and the height is

$$\Delta h = h(t + \varepsilon) - h(t) - v\varepsilon \quad (33)$$

The height $h(t)$ was found previously as

$$h(t) = \frac{c^2 t^2 - z_0^2}{2[R_0 + z_0]}$$

that plus the two right triangles T1 and T2 can be solved for $h(t + \varepsilon)$:

$$\text{T1: } c^2[t + \varepsilon]^2 = R^2 + [z_0 - v\varepsilon + h(t + \varepsilon)]^2 \quad (34)$$

$$\text{T2: } [R_0 + v\varepsilon]^2 = R^2 + [R_0 + v\varepsilon - h(t + \varepsilon)]^2 \quad (35)$$

$$\text{T1 - T2: } c^2[t + \varepsilon]^2 - [R_0 + v\varepsilon]^2 = [z_0 - v\varepsilon + h(t + \varepsilon)]^2 - [R_0 + v\varepsilon - h(t + \varepsilon)]^2 \quad (36)$$

Expanding the terms on the left side of T1–T2 gets:

$$\begin{aligned} c^2[t + \varepsilon]^2 - [R_0 + v\varepsilon]^2 &= c^2t^2 + 2c^2t\varepsilon + c^2\varepsilon^2 - R_0^2 - 2R_0v\varepsilon - v^2\varepsilon^2 \\ &= c^2t^2 + 2c^2t\varepsilon - R_0^2 - 2R_0v\varepsilon \text{ as } \varepsilon \rightarrow 0 \end{aligned} \quad (37)$$

Expanding the two terms on the right side of T1–T2 yields for the first term:

$$[z_0 - v\varepsilon + h(t + \varepsilon)]^2 = h^2(t + \varepsilon) + z_0^2 + v^2\varepsilon^2 + 2h(t + \varepsilon)z_0 - 2z_0v\varepsilon - 2h(t + \varepsilon)v\varepsilon \quad (38)$$

and yields for the second term:

$$[R_0 + v\varepsilon - h(t + \varepsilon)]^2 = R_0^2 + v^2\varepsilon^2 + h^2(t + \varepsilon) + 2R_0v\varepsilon - 2v\varepsilon h(t + \varepsilon) - 2R_0h(t + \varepsilon) \quad (39)$$

Simplifying and subtracting the two terms, yields for the right side of T1–T2:

$$z_0^2 - R_0^2 + 2h(t + \varepsilon)z_0 - 2R_0v\varepsilon - 2z_0v\varepsilon + 2R_0h(t + \varepsilon) \quad (40)$$

Equating the left side to the right side of T1–T2 produces

$$c^2t^2 + 2c^2t\varepsilon - R_0^2 - 2R_0v\varepsilon = z_0^2 - R_0^2 + 2h(t + \varepsilon)z_0 - 2R_0v\varepsilon - 2z_0v\varepsilon + 2R_0h(t + \varepsilon) \quad (41)$$

$$\begin{aligned} c^2t^2 + 2c^2t\varepsilon &= z_0^2 + 2h(t + \varepsilon)z_0 - 2z_0v\varepsilon + 2R_0h(t + \varepsilon) \\ &= z_0^2 - 2z_0v\varepsilon + 2h(t + \varepsilon)[R_0 + z_0] \end{aligned} \quad (42)$$

Solving for $h(t + \varepsilon)$ results in:

$$h(t + \varepsilon) = \frac{c^2t^2 - z_0^2 + 2c^2t\varepsilon - 2z_0v\varepsilon}{2[R_0 + z_0]} \quad (43)$$

Then

$$\Delta h = h(t + \varepsilon) - h(t) - v\varepsilon = \frac{2c^2t\varepsilon - 2z_0v\varepsilon}{2[R_0 + z_0]} - v\varepsilon = \varepsilon \left[\frac{c^2t - z_0v}{[R_0 + z_0]} - v \right] \quad (44)$$

and squaring both sides and taking a limit gets

$$\Delta h^2 = \varepsilon^2 \left[\frac{c^2t - z_0v}{[R_0 + z_0]} - v \right]^2 = 0, \text{ as } \varepsilon \rightarrow 0 \quad (45)$$

So the frustum area as $\varepsilon \rightarrow 0$ is

$$A_{sv} = \pi[R_1 + R_2]d = \pi[R_1 + R_2]\sqrt{\Delta h^2 + (R_2 - R_1)^2} = \pi[R_1 + R_2][R_2 - R_1] = \pi[R_2^2 - R_1^2] \quad (46)$$

which is also the area, A_p , of an annulus with radii R_1 and R_2 .

This annulus area $A_p = 2\pi c^2t\varepsilon$ was previously computed for the stationary planar source and was also used in determining the area of the nonexpanding spherical source, $A_s = [R_0/[R_0 + z_0]]2\pi c^2t\varepsilon$. Consequently the previously developed results for the non expanding sphere can be used if the effects of the motion of the expanding spherical source on its wavelet point sources' initial impulses are included:

At the wavelet point source location as $z_0 \rightarrow 0$ the motion due to the sphere's expanding radius is the same as the motion of the moving planar source if spherical spreading is disregarded. Then the method used for the moving planar source will be used here to implement the effects of motion on the wavelet point sources initial impulses due to the sphere's radius expanding with a speed v :

As with the non expanding sphere, the wavelet point sources at points N and F of the near and far surfaces of the expanding sphere are the sole contributors to the wave front along the radial line. The wavelet point sources at points N and F on the near and far surfaces have relative speeds $+v$ and $-v$ with respect to their originating surfaces. (The far surface is the source for a backward propagating wave front and so has relative speed $-v$.)

Including the moving planar source's source speed effect in the equation for the wave front from a nonexpanding spherical source observed at P as a function of time yields:

$$f(z_0, t) = \frac{R_0}{2[R_0 + z_0]} \left[\frac{v + c}{c} \delta\left(t - \frac{z_0}{c}\right) - \frac{-v + c}{c} \delta\left(t - \frac{z_0 + 2R_0}{c}\right) \right] \quad (47)$$

for the forward propagating spherical wave front from the near surface combined with the backward propagating spherical wave front from the far surface of the expanding spherical source.

When source speed equals zero, $v = 0$, the result is the same as for the non expanding spherical source.

When source speed equals the propagation speed, $v = c$, the equation becomes

$$f(z_0, t) = \frac{R_0}{2[R_0 + z_0]} \left[\frac{c + c}{c} \delta\left(t - \frac{z_0}{c}\right) - \frac{-c + c}{c} \delta\left(t - \frac{z_0 + 2R_0}{c}\right) \right] = \frac{R_0}{R_0 + z_0} \delta\left(t - \frac{z_0}{c}\right) \quad (48)$$

So when $v = c$ the backward wave front is eliminated and $f(z_0, t)$ yields an expanding forward propagating impulsive spherical wave front, which is an impulse when seen at P as a function of time. (Note the absence of

the $\frac{1}{2}$ factor). Since this holds for every radial line, this is the new impulsive spherical progressing Huygens' wave front which has neither a wake nor a backward wave.

The spherical spreading factor $\frac{R_0}{R_0+z_0}$ differentiates this result from the result from the moving planar source and accounts for the radial lines representing infinitesimal solid angles. As $R_0 \rightarrow \infty$ this equation becomes the same as for the planar source moving with speed equal to c .

Huygens' Principle as depicted in his figures. It was shown that for the spherical and planar source, when the excitation is $\delta(t)$, the resulting wave field ϕ when differentiated disappears except at the locations of the propagating wave fronts f . The following shows how the intermediate step of differentiation can be eliminated.

Let L represent the linear wave propagation operator which transforms the excitation on either the plane or sphere into the wave field ϕ . If $\delta(t)$ is the excitation, the resulting wave field due to the wave propagation process is

$$\phi(z, t) = L(\delta(t)) \quad (49)$$

which is the response of L to an impulse. As a function of time in the case of the plane this wave field is a step function and in the case of the sphere it is a rectangular pulse.

Because L is linear (LTI) the derivative can be taken of both ϕ and of $\delta(t)$ ⁴² yielding

$$\frac{\partial}{\partial t} \phi(z, t) = L\left(\frac{d}{dt} \delta(t)\right) = L(\delta'(t)) \quad (50)$$

But the derivative of the wave field ϕ is the wave front f , so

$$f(z, t) = \frac{\partial}{\partial t} \phi(z, t) = L(\delta'(t)) \quad (51)$$

As a result, if the doublet $\delta'(t)$ is used as the initial excitation instead of the singlet $\delta(t)$ for the wavelets originating from the source, the Huygens' wavelets when combined disappear except where they contact their common tangent surfaces. Also when the sources' motion equals the wave propagation speed the backward wave corresponding to the backward tangent surface disappears. The remaining forward tangent surface is the new propagating wave front, obtained without the intermediate step of differentiation.

In Huygens' figures the wavelets then would represent doublets $[4\pi r]^{-1} \delta'(t - r/c)$. However, notice that the resulting new wave fronts are of the form δ not δ' so they must be differentiated before being used as the initial excitation in the next iteration of wave propagation.

Because ϕ is the response to an impulse, the analysis can be applied to other forms of excitation, say $g(t)$, by using convolution: the convolution

$$L(g) = g * L(\delta) = g * \phi \quad (52)$$

gives the response of L to the excitation g .

Summary

Huygens' geometrical construction has been shown to be literally correct as depicted in his figures (see Fig. 1a) when his wavelets represent the first derivative of Dirac delta distributions and the source motion is equal to the propagation speed. For the sphere and the infinite plane there is no wake and there is no backward wave, thus solving the two most long standing problems (1678).

An intuitive, relatively simple method to analyze wave propagation has been provided which provides a clear view of the propagation process starting with the very first wave propagation analysis, Huygens' Principle. This will facilitate the teaching of wave physics and provide a new model for wave propagation.

Received: 15 July 2021; Accepted: 14 September 2021

Published online: 12 October 2021

References

- Huygens, C. *Treatise on Light* (Macmillan and Co., 1912).
- Hadamard, J. *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Dover, 1952).
- Baker, B. B. & Copson, E. T. *The Mathematical Theory of Huygens' Principle* (Chelsea Publishing Company, 1987).
- Miller, D. A. B. Huygens wave propagation principle corrected. *Opt. Lett.* **16**, 1370–1372 (1991).
- Daniels, J. M. An explanation of a difficulty with Huygens secondary wavelets. *Can. J. Phys.* **74**, 236–239 (1996).
- Blok, H. P. et al. (eds) *Huygens' Principle, 1690–1990: Theory and Applications: Proceedings of an International Symposium, The Hague/Scheveningen, November 19–22, 1990* (Elsevier, 1992).
- Gunther, P. *Huygens' Principle And Hyperbolic Equations* (Academic Press, 1988).
- Rumsey, V. H. *Huygens' Principle as an Exact Physical Concept* (Berkeley, Calif: University of California, Division of Electrical Engineering, Electronics Research Laboratory, IER report ser. 60, no. 128 1954)
- Madrid, J. A. A geometric approach to time evolving wavefronts. *Geophys. J. Int.* **172**, 1117–1122 (2008).
- Luis, A. Complementary Huygens principle for nongeometrical optics. *Eur. J. Phys.* **28**, 231–240 (2007).
- Enders, P. The Ewald–Oseen extinction theorem in light of Huygens' principle. *Electron. J. Theor. Phys.* **41**, 127–136 (2011).
- Hansen, T. & Kaiser, G. Generalized Huygens principle with pulsed beam wavelets. *J. Phys. A Math. Theoret.* **42**, 475403 (2009).
- Kaiser, G. Huygens' principle in classical electrodynamics: A distributional approach. *Adv. Appl. Clifford Algebras* **22**, 703–720 (2009).
- Mende, F. F. Physical substantiation of Huygens principle and the reciprocity theorem. *Am. J. Electr. Electron. Eng.* **2**(6), 165–170 (2014).

15. Shul'ga, N. F. & Dobrovolskij, S. N. Huygens principle in the problem of relativistic electrons transitions radiation in thin metal target distributed transverse size. *Izvestiya Akad. Nauk Rossijskaya Akad. Nauk Seriya Fizicheskaya* **11**, 2174–2178 (2000).
16. Ivanov, L. A. Huygens' principle in an even-dimensional space for some equations with singularities. *Ukrain. Math. J.* **31**, 433–435 (1979).
17. Yura, K. O. Huygens' principle for the Dirichlet boundary value problem for the wave equation. *Sci. Math. Jpn.* **3**, 419–426 (2004).
18. Werby, M. F. A one way coupled mode solution, Huygens' principle and mode coupling coefficients. *J. Acoust. Soc. Am.* **106**, 2225 (1999).
19. Markle, C. D. & Cowsik, R. General approach to Casimir force problems based on local reflection amplitudes and Huygens' principle. *Phys. Rev. A* **85**, 052516 (2012).
20. Illert, C. R. Origins of linguistic zonation in the Australian Alps part 1–Huygens' principle. *J. Appl. Stat.* **6**, 625–659 (2005).
21. Ueda, M. Extension of the Huygens–Fresnel principle to a virtual space. *J. Acoust. Soc. Am.* **96**, 3226 (1994).
22. Censor, D. The Huygens' principle and Doppler effect in the presence of time-dependent surfaces. *J. Electromagn. Waves Appl.* **3**, 635–649 (1989).
23. Ghavami, N., Smith, P. P., Tiberi, G., Edwards, D. & Craddock, I. Non-iterative beamforming based on Huygens principle for multistatic ultrawide band radar: Application to breast imaging. *IET Microw. Antennas Propag.* **9**, 1233–1240 (2015).
24. Tabin, J. New aspects of Huygens principle. *Appl. Phys. B* **31**, 225–228 (1983).
25. Solomon, J. M. Huygens' principle for a class of singular Cauchy problems. *J. Differ. Equ.* **10**, 219–239 (1971).
26. Carminati, J. & McLennaghan, R. G. The validity of Huygens' principle for the conformally invariant scalar wave equation, Maxwell's equations and Weyl's neutrino equation on Petrov type D and type III space-times. *Phys. Lett. A* **118**, 322–324 (1986).
27. Eleuch, H. & Rostovtsev, Y. V. Analytical solution for 3D stationary Schrödinger equation: Implementation of Huygens' principle for matter waves. *J. Mod. Opt.* **57**, 1877–1881 (2010).
28. Chauvat, D., Emile, O., Brunel, M. & Le Floch, A. Huygens' principle and Young's experiment in the propagation of light beams. *Am. J. Phys.* **71**, 1196–1198 (2003).
29. Ooguri, H. Spectrum of Hawking radiation and the Huygens principle. *Phys. Rev. D* **33**, 3573 (1986).
30. Rumsey, V. Some new forms of Huygens' principle. *IRE Trans. Antennas Propag.* **7**, 103–116 (1959).
31. Darling, B. T. Continuation of solutions and the retarded electromagnetic fields in the mathematical formulation of Huygens' principle. *J. Mod. Opt.* **31**, 97–105 (1984).
32. Loewenthal, D. Huygens principle versus exploding reflector? theoretical and numerical aspects. *Explor. Geophys.* **27**, 183–186 (1996).
33. Piaggio, H. The mathematical theory of Huygens' principle. *Nature* **145**, 531–532 (1940).
34. Zemanian, A. H. *Distribution Theory And Transform Analysis* (McGraw-Hill, 1965).
35. Friedlander, F. G. *Introduction to the Theory of Distributions* (Cambridge University Press, 1982).
36. Schwartz, L. *Mathematics For The Physical Sciences* (Dover Publications, 2008).
37. Gelfand, I. M. & Shilov, G. E. *Generalized Functions, Volume 1, Properties and Operations* (Academic Press, 1964).
38. Farassat, F. *Introduction to Generalized Functions with Applications in Aerodynamics and Aeroacoustics* Paper 3428 (Langley Research Center, 1994).
39. Feng, S. & Winful, H. G. Physical origin of the Gouy phase shift. *Opt. Lett.* **26**, 485–487 (2001).
40. Boyd, R. W. Intuitive explanation of the phase anomaly of focused light beams. *J. Opt. Soc. Am.* **70**, 877–880 (1980).
41. Ruffin, A. B., Rudd, J. V., Whitaker, J. F., Feng, F. & Winful, H. G. Direct observation of the gouy phase shift with single cycle terahertz pulses. *Phys. Rev. Lett.* **83**, 3410–3413 (1999).
42. Smith, S. *The Scientist's and Engineer's Guide to Digital Signal Processing* 134, 135 (California Technical Publishing, 1999).

Author contributions

F.A. did the analysis, drew the figures, and wrote the manuscript.

Competing interests

The author declares no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-99049-7>.

Correspondence and requests for materials should be addressed to F.L.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021