Wave propagation: Odd is better, but three is best

2. Propagation in spaces of different dimensions

V. Balakrishnan
Department of Physics, Indian Institute of Technology Madras
Chennai 600 036, India

In Part of 1 of this article, we have shown that the basic solution to the wave equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) u(\mathbf{r}, t) = \delta^{(D)}(\mathbf{r} - \mathbf{r}_0)\,\delta(t - t_0) \tag{1}$$

that vanishes as $r \to \infty$ is given by

$$u^{(D)}(\mathbf{R}, \tau) = c\,\theta(\tau) \int \frac{d^D\mathbf{k}}{(2\pi)^D}\,\frac{\sin\,c\tau k}{k}\,e^{i\mathbf{k}\cdot\mathbf{R}}\,, \tag{2}$$

where D is the number of spatial dimensions, $\mathbf{R} \equiv \mathbf{r} - \mathbf{r}_0$ and $\tau \equiv t - t_0$. We now simplify and analyse the solution for different values of D.

The case $D = 1$

The case of a single spatial dimension is somewhat distinct from the others, and simpler too. Let us dispose of this case first.

Recall that the symbol k in the factor $(\sin\,c\tau k)/k$ in Eq. (2) stands for $|\mathbf{k}|$; in the case $D = 1$, therefore, we should remember to write $|k|$ instead of just k in this factor. Further, $\mathbf{k}\cdot\mathbf{R}$ is just kX in this case, where $X = x - x_0$. Therefore

$$u^{(1)}(X, \tau) = c\;\theta(\tau) \int_{-\infty}^{\infty} \frac{dk}{2\pi}\,\frac{\sin\,c\tau|k|}{|k|}\,e^{ikX} = c\;\theta(\tau) \int_{-\infty}^{\infty} \frac{dk}{2\pi}\,\frac{\sin\,c\tau k}{k}\,e^{ikX}\,. \tag{3}$$

It is immediately evident from this expression that $u^{(1)}(-X, \tau) = u^{(1)}(X, \tau)$, i. e., that $u^{(1)}(X, \tau)$ is in fact a function of $|X|$. Using $e^{ikX} = \cos\,kX + i\,\sin\,kX$, we see that the contribution from the $\sin\,kX$ term vanishes because the integrand is an odd function of k. Thus

$$\begin{aligned} u^{(1)}(X, \tau) &= c\,\theta(\tau) \int_0^{\infty} \frac{dk}{\pi k}\,\sin\,(c\tau k)\,\cos\,kX \\ &= c\,\theta(\tau) \int_0^{\infty} \frac{dk}{2\pi k}\Big(\sin\,(c\tau + X)k + \sin\,(c\tau - X)k\Big) \\ &= \frac{c}{4}\,\theta(\tau)\,\Big(\varepsilon(c\tau + X) + \varepsilon(c\tau - X)\Big)\,. \end{aligned} \tag{4}$$

In the last equation, we have used the well-known fact that $\int_0^{\infty} dk\,(\sin\,bk)/k = (\pi/2)\varepsilon(b)$ for any real number b ; here, the symbol $\varepsilon(b) = \theta(b) - \theta(-b) = +1$ for

$b > 0$, and $\varepsilon(b) = -1$ for $b < 0$. Simplifying the final expression in Eq. (4), we find

$$u^{(1)}(X, \tau) = \frac{c}{2}\,\theta(\tau)\,\theta(c\tau - |X|)\,. \tag{5}$$

The second step function ensures that the signal does not reach any point x until time $t_0 + |x - x_0|/c$, as required by causality. The presence of this step function makes the other step function, $\theta(\tau)$, redundant from a physical point of view. However, it is present in the formal mathematical solution for the quantity $u^{(1)}(X, \tau)$.

But there is another aspect of the solution which is noteworthy. Although an observer at an arbitrary point x starts receiving the signal at time $t_0 + |x - x_0|/c$, he does not receive a *pulsed* signal, even though the sender sent out such a signal. In fact, the signal received *persists* thereafter for all time, without diminishing in strength! This last feature is peculiar to $D = 1$. Let us see what happens in higher dimensions.

The case $D = 2$

Before we discuss the nature of the solution for $D \geq 2$, we must note an important feature of $u^{(D)}(\mathbf{R}, \tau)$.

The expression in Eq. (2) is a *scalar*: by this we mean that it is unchanged under rotations of the spatial coordinate axes about the origin. This remains true for all integer values of $D \geq 2$.

This assertion may seem to be more-or-less obvious, because $\mathbf{k} \cdot \mathbf{R}$ is after all a scalar product of two D-dimensional vectors. But it must be proved rigorously, which requires a bit of work. We will not do so here, in the interests of brevity, but merely point out that two factors play a role in such a proof. First, the region of integration in Eq. (2) is all of $\mathbf{k}$-space, and this is invariant under rotations of the coordinate axes in that space. Second, the volume element $d^{(D)}\mathbf{k}$ is also similarly unchanged under rotations of the axes.

As a result of this rotational invariance, $u^{(D)}(\mathbf{R}, \tau)$ is actually a function of R and τ (where $R \equiv |\mathbf{R}|$, as already defined). The consequence of this is that we can choose the orientation of the axes in $\mathbf{k}$-space according to our convenience, without affecting the result.

Turning now to the $D = 2$ case, it is evidently most convenient to work in plane polar coordinates, choosing the k_1-axis along the vector $\mathbf{R}$. Then

$$\begin{aligned} u^{(2)}(R, \tau) &= c\,\theta(\tau) \int_0^\infty \frac{k\,dk}{(2\pi)^2}\,\frac{\sin c\tau k}{k} \int_0^{2\pi} d\varphi\, e^{ikR\cos\varphi} \\ &= c\,\theta(\tau) \int_0^\infty \frac{dk}{2\pi}\,\sin(c\tau k)\,J_0(kR)\,, \end{aligned} \tag{6}$$

where $J_0(kR)$ is the Bessel function of order 0. The final integral over k is again a known integral, equal to $(c^2\tau^2 - R^2)^{-1/2}$ provided $c^2\tau^2 > R^2$, and zero otherwise.

Since we are concerned here with the physical region in which both τ and R are non-negative, our solution reads

$$u^{(2)}(R, \tau) = \frac{c\,\theta(\tau)}{2\pi} \frac{\theta(c\tau - R)}{\sqrt{c^2\tau^2 - R^2}}\,. \tag{7}$$

The signal thus reaches any point $\mathbf{r}$ only at time $t_0 + |\mathbf{r} - \mathbf{r}_0|/c$, in accordance with causality and the finite velocity of propagation of the disturbance. But once again, the signal is no longer a sharply pulsed one: it persists for all $t > t_0 + |\mathbf{r} - \mathbf{r}_0|/c$, although its strength slowly decays as t increases, like $1/t$ at very long times.

The case $D = 3$

Something entirely different happens in three-dimensional space. We have

$$u^{(3)}(R, \tau) = c\,\theta(\tau) \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{\sin\, c\tau k}{k}\, e^{i\mathbf{k}\cdot\mathbf{R}}\,. \tag{8}$$

Rotational invariance is now exploited: we use spherical polar coordinates (k, θ, φ) in $\mathbf{k}$-space, and, moreover, *choose the polar axis along the vector* $\mathbf{R}$. This immediately enables us to carry out the integration over the azimuthal angle φ, obtaining a factor 2π. It is useful to write out the subsequent steps in this instance, because they (or their variants) appear in more than one context in physical applications.

$$\begin{aligned}
u^{(3)}(R, \tau) &= \frac{c\,\theta(\tau)}{(2\pi)^2} \int_0^\infty dk\, k^2\, \frac{\sin\, c\tau k}{k} \int_{-1}^{1} d(\cos\,\theta)\, e^{ikR\,\cos\,\theta} \\
&= \frac{2c\,\theta(\tau)}{(2\pi)^2 R} \int_0^\infty dk\, \sin\,(c\tau k)\, \sin\,(kR) \\
&= \frac{c\,\theta(\tau)}{(2\pi)^2 R} \int_0^\infty dk\, \Big(\cos\,(c\tau - R)k - \cos\,(c\tau + R)k \Big) \\
&= \frac{c\,\theta(\tau)}{2(2\pi)^2 R} \int_{-\infty}^\infty dk\, \Big(\cos\,(c\tau - R)k - \cos\,(c\tau + R)k \Big) \\
&= \frac{c\,\theta(\tau)}{2(2\pi)^2 R}\, \mathrm{Re} \int_{-\infty}^\infty dk\, \Big(e^{i(c\tau - R)k} - e^{i(c\tau + R)k} \Big) \\
&= \frac{c\,\theta(\tau)}{4\pi R}\, \mathrm{Re}\, \Big(\delta(c\tau - R) - \delta(c\tau + R) \Big).
\end{aligned} \tag{9}$$

But the delta functions are real quantities. And once again, we are interested in the region in which both τ and R are non-negative. The solution therefore reduces to

$$u^{(3)}(R, \tau) = \frac{c\,\theta(\tau)\,\delta(c\tau - R)}{4\pi R} = \frac{\theta(\tau)\,\delta(\tau - R/c)}{4\pi R}. \tag{10}$$

Thus, almost miraculously, the signal is also a delta function pulse that reaches (and passes) an observer at any point $\mathbf{r}$ at precisely the instant $t_0 + |\mathbf{r} - \mathbf{r}_0|/c$.

There is no after-effect that lingers on, in stark contrast to the situation in $D = 1$ and $D = 2$.

The amplitude of the pulse drops with distance like $1/R$, exactly the way the Coulomb potential does. In fact, this is yet another unique feature of the solution in $D = 3$. Formally, if the limit $c \to \infty$ is taken in Eq. (1), the wave operator reduces to the negative of the Laplacian operator. We might therefore expect the solution for $u(\mathbf{r}, t)$ to reduce to the corresponding Green function for $-\nabla^2$. In three dimensions, this is precisely $1/(4\pi R)$. This fact is very familiar to us from electrostatics. The potential $\phi(\mathbf{r})$ due to a point charge q located at $\mathbf{r}_0$ satisfies the equation $-\nabla^2 \phi(\mathbf{r}) = \rho(\mathbf{r})/\epsilon_0 = (q/\epsilon_0)\, \delta^{(3)}(\mathbf{r} - \mathbf{r}_0)$. With the boundary condition $\phi \to 0$ as $r \to \infty$, the solution to this equation is just Coulomb's Law, namely, $\phi(\mathbf{r}) = -q/(4\pi\epsilon_0 R)$ where $R = |\mathbf{r} - \mathbf{r}_0|$. This reduction of the solution of the inhomogeneous wave equation to that of Poisson's equation in the limit $c \to \infty$ does *not* occur in $D = 1$ or $D = 2$.

Dimensions $D > 3$

Now that we have appreciated a very important feature of three-dimensional space that is absent in one- and two-dimensional spaces, it is natural to ask if this feature is unique to $D = 3$. Surprisingly, it is not: the propagation of sharp signals is possible in all *odd*-dimensional spaces with $D \geq 3$, while it fails for all *even* values of D. In other words, the signal received at any point $\mathbf{r}$ lingers on for all $t > t_0 + |\mathbf{r} - \mathbf{r}_0|/c$ in $D = 2, 4, \cdots$, while it is sharply pulsed, arriving and passing on at time $t_0 + |\mathbf{r} - \mathbf{r}_0|/c$ with no after-effect, in $D = 3, 5, \cdots$. There is, however, one feature that is absolutely *unique* to $D = 3$: this is the only case in which the original δ-function pulse is transmitted without any *distortion*, namely, as a δ-function pulse.

One way to establish these results is to start with Eq. (2), and to use hyper-spherical coordinates in D dimensions. Then $\mathbf{k} = (k, \theta_1, \theta_2, \ldots, \theta_{D-2}, \varphi)$ where $0 \leq k < \infty\,,\ 0 \leq \theta_i \leq \pi\,,\ 0 \leq \varphi < 2\pi$. Once again, we may choose the k_1 axis to lie along the vector $\mathbf{R}$, which permits us to carry out the integrations over $\theta_2, \ldots, \theta_{D-2}$ and φ. The result is

$$u^{(D)}(R, \tau) = (\text{const.})\, \theta(\tau) \int_0^\infty dk\, k^{D-2}\, \sin(c\tau k) \int_0^\pi d\theta_1\, (\sin\, \theta_1)^{D-2}\, e^{ikR\, \cos\, \theta_1} \tag{11}$$

where the constant depends on D. Clearly, this is a laborious method of finding $u^{(D)}(R, \tau)$, especially as the integrations over θ_1 and k have yet to be carried out.

There is a more elegant and powerful way to solve the problem. This is based on the *relativistic invariance* of the wave operator and the solution sought. A detailed account of this would take us too far afield. We therefore restrict ourselves to a short description of this approach, to get some feel for the underlying "mechanism" responsible for the basic difference between the cases of even and odd D. Our discussion will not be fully rigorous, as we shall not pay attention to certain technical details that warrant a more careful examination.

The operator $(1/c^2)\,\partial^2/\partial t^2 - \nabla^2$ can be verified to be unchanged in form ("invariant") under Lorentz transformations in $(D+1)$-dimensional space-time. As a consequence of this invariance, the specific solution we seek can also be shown to be Lorentz-invariant. In the present context, this means that we can always evaluate the integrals involved in Eq. (2) by first transforming to an inertial frame in which the four-vector $(c\tau, \mathbf{R})$ has only a time-like component, i.e., it is of the form $(c\tau\,', \mathbf{0})$, where $c^2\tau^2 - R^2 = c^2\tau\,'^2$. [This can only be done for a so-called *time-like* four vector, i.e., one for which $c^2\tau^2 - R^2 > 0$. It cannot be done for a light-like four-vector ($c^2\tau^2 - R^2 = 0$) or a space-like four-vector ($c^2\tau^2 - R^2 < 0$). This is the technical point we slur over, with the remark that our conclusions will not be affected by it.] After the integrals required are evaluated, we can transform back to the original frame by replacing $c\tau\,'$ with $(c^2\tau^2 - R^2)^{1/2}$. We must also mention that $\tau > 0$ implies $\tau\,' > 0$, because the sign of the time component of a four-vector remains unchanged under the set of Lorentz transformations with which we are concerned.[1] Denoting the corresponding signal by $u^{(D)}(\tau\,')$, we have

$$u^{(D)}(\tau\,') = c\,\theta(\tau\,') \int \frac{d^D\mathbf{k}}{(2\pi)^D}\, \frac{\sin\, c\tau\,'k}{k} = (\text{const.})\,\theta(\tau\,') \int_0^\infty dk\, k^{D-2}\, \sin\,(c\tau\,'k)\,, \tag{12}$$

on carrying out all the angular integrals in D-dimensional space. The constant on the RHS in the last equation depends on D. This representation shows us, in very clear fashion, how the cases of odd and even D differ from each other. When D is odd, the integrand is an even function of k, and hence the integral can be converted to one that runs from $-\infty$ to ∞. The result can then be shown to be essentially a derivative of a certain order of the delta function $\delta(c^2\tau\,'^2)$, i. e., a sharply pulsed signal. (The order of the derivative increases with D.) On the other hand, when D is even, this cannot be done, and the integral leads to an *extended* function of $c^2\tau\,'^2$. This dissection lays bare the precise mathematical distinction that lies at the root of the physical differences in signal propagation in odd and even dimensional spaces, respectively. In fact, the formal solution for $u^{(D)}(\tau\,')$ can be shown to be essentially the derivative of order $(D-3)/2$ of $\delta(c^2\tau\,'^2)$ in *all* cases. When D is even, this is a so-called *fractional derivative*, which is a non-local object — in physical terms, an extended function.

The form of the result in Eq. (12) suggests even more. Since the second derivative of the sine function is again a sine function (apart from a minus sign), it follows that the solution in $(D+2)$ spatial dimensions can be obtained from that in D space dimensions by a simple trick. We find

$$u^{(D+2)}(\tau\,') = -\frac{1}{2\pi c^2 D}\, \frac{\partial^2 u^{(D)}(\tau\,')}{\partial \tau\,'^2}\,. \tag{13}$$

[1]Once again, this is only true for a time-like or light-like four-vector, but not a space-like one.

This shows how the solutions in $D = 5, 7, \ldots$ can be generated from that in $D = 3$, while those in $D = 4, 6, \ldots$ can be generated from that in $D = 2$. The detailed working out of these solutions is left to the interested reader.

A final remark, before we pass on to more general considerations. How widely applicable are the conclusions at which we have arrived? Basically, there are two important additional aspects of wave or signal propagation that can be adjusted so as to modify the basic result. The first is *dispersion*. Sinusoidal waves of different wavelengths will, in general, propagate with different speeds in a medium. The precise manner in which the frequency and wavelength of waves in a medium are related to each other is called a dispersion relation. Such relations can be quite complicated. The second aspect is *nonlinearity*. The simple wave equation we have used, Eq. (1), is *linear* in u. On the other hand, physical situations often call for nonlinear equations. The interplay between dispersion and nonlinearity can be extremely intricate and interesting, and a vast variety of new phenomena can arise as a result. Among these are the so-called solitary waves and propagating solitons, which represent very robust pulsed disturbances.

General remarks on dimensionality

Is there anything else special about *three*-dimensional space that is not shared by a space of any other dimensionality? Again, answers can be given at many levels. An important observation is that it is only in $D = 3$ that the cross product of two vectors is again a vector. For, it is only in $D = 3$ that the number of mutually perpendicular planes spanning the space is equal to the number of Cartesian coordinate axes, 3 being the only nonzero solution of the equation ${}^{D}C_2 \equiv D(D-1)/2 = D$. Though these statements appear to be simple-minded, they have profound consequences.

At a slightly more sophisticated level, we may make the following general, if rather loose, statement: one- or two-dimensional space is, in some sense, too "simple" for anything too complicated to be possible; on the other hand, four- or higher-dimensional space is again too "roomy" for anything very complicated to occur. (In even more loose terms, this "roominess" permits the *undoing* of complications like knots, for instance.) This leaves $D = 3$ as the most "interesting" number of dimensions. Once again, we refrain from further elaboration.

Finally, we must recognize that our sensory organs and the information processing hardware and software in our brains are designed so specifically for (3+1)-dimensional space-time, that we literally take this dimensionality to be a fundamental and "self-evident" fact of nature. In actuality, however, there are very deep unanswered questions about the nature of space and time. These questions are connected to questions about quantum mechanics, general relativity and the origin of the universe. We do not know for sure whether, at the very smallest time scales and length scales, the number of space dimensions is three or more; or whether space-time coordinates must be supplemented with certain other kinds of variables to specify a point in the "true" arena in which phenomena occur; or even whether space-time is ultimately continuous or discrete ("granular"). One thing does appear to be fairly certain, though: *it is very probable that, sooner or later, our long-standing ideas and preconceptions about the nature of space and time will have to be revised significantly at the most fundamental level.*

I thank S. Govindarajan, S. Lakshmi Bala and N. Mukunda for a critical reading of the manuscript and for helpful comments.

Suggested reading

G. Barton, *Elements of Green's Functions and Propagation — Potentials, Diffusion and Waves*, Clarendon Press, Oxford, 1989.

R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2, Interscience, New York, 1962.

P. R. Garabedian, *Partial Differential Equations*, Wiley, New York, 1964.

G. B. Whitham, *Linear and Nonlinear Waves*, Wiley, New York, 1999.

Box

The French mathematician, natural philosopher and encyclopaedist Jean le Rond d'Alembert (1717-1783) was among the first to understand the significance of, and study in some detail, several important differential equations of mathematical physics. Among other results, he showed that the general solution of the one-dimensional wave equation $\partial^2 u/\partial t^2 - c^2\,\partial^2 u/\partial x^2 = 0$ is of the form $u(x,t) = f_1(x+ct) + f_2(x-ct)$. This corresponds to the superposition of two different waveforms or pulses moving, respectively, to the left and right with speed c. It is the forerunner of *the method of characteristics* for a class of partial differential equations. D'Alembert's name is associated with many other discoveries as well, such as d'Alembert's Principle in Mechanics, d'Alembert's scheme in games of chance such as roulette, and d'Alembert's paradox: he showed that, in the streamlined, irrotational flow of a non-viscous fluid past a solid obstacle, the net drag force on the solid *vanishes*, contrary to what one would guess off-hand.

D'Alembert (along with the philosopher Denis Diderot) spent a good deal of time and effort on a massive project, the great French Encyclopaedia. (By the way, the popular hilarious story about the great mathematician Euler confounding Diderot with his spoof of a "mathematical proof" of the existence of God appears to be — sadly enough — without foundation.) D'Alembert seems to have been a ‘straight shooter’; according to W. W. Rouse Ball, “d’Alembert’s style is brilliant but not polished, and faithfully reflects his character, which was bold, honest and frank. $\cdots$ with his dislike of sycophants and bores it is not surprising that during his life he had more enemies than friends.”