Qeios

Inadequacies of Sommerfeld's Front Velocity Definition

Steffen Kühn
steffen.kuehn@aurinovo.de

July 13, 2024

***Abstract*—Current practice defines the front velocity of a signal as the limit of the phase velocity for infinitely high frequency. However, the present article provides evidence that the propagation velocities of signal fronts for input signals of nonzero temporal duration result from the phase velocities in the low-frequency range. In conclusion, although the impulse response propagates at the phase velocity for infinitely high frequency, this is shown not to be true for the step response. The article further clarifies that the front or information velocity should be defined as the fastest phase velocity of a transmission medium.**

***Index Terms*—front velocity, phase velocity, signal velocity**

I. Introduction

The front velocity is defined in the scientific literature as the limit of the phase velocity $v_p(\omega \to \infty) = c$ (e.g., [1]–[4]). However, this article mathematically demonstrates by means of a causal counterexample that the current definition of the front velocity is incorrect. In fact, only the speed of the impulse response of a transmission line is limited by c. The speed of the step response, in contrast, is dominated by the phase velocities in the low-frequency range.

At a first glance, this aspect seems to be a contradiction, because the step response is the convolution of the impulse response with the unit step function. However, this article shows that no paradox exists, and the step response could move faster than c if the transmission line had phase velocities faster than c.

The definition of the front velocity as the limit of the phase velocity originates from a paper published in 1914 by the well-known theoretical physicist Arnold Sommerfeld [5]. His mathematical reasoning is essentially based on the discontinuity in a sine signal when the signal is turned on. This discontinuity consists of arbitrarily high frequencies which propagate with phase velocity $v_p(\omega \to \infty) = c$. On this basis, he concluded that a front, i.e., a sudden change in the signal level, likewise cannot propagate faster than at a speed of c.

These considerations continue to represent the foundation of what is considered the current state of the art [2], [6]. Often this definition is used to shortcut a discussion or to argue that a transmission of information at a speed faster than light in vacuum is principally impossible. This might be the case, but the reason is probably that there are no transmission media that have phase velocities that exceed the speed of light.

This seems to be in contradiction with the term *superluminal phase velocity* which can occasionally be found in the literature [3], [7]–[9]. We note that this refers to phase velocities of partially standing waves, i.e., waves in which a part of the wave is moving in one direction and another part is moving in the opposite direction. Such partially standing waves appear when electromagnetic waves are reflected from surfaces or molecules in the transmission medium. The phase velocities of such partially standing waves are not one-way phase velocities and do not represent true phase velocities.

II. Starting point

To make it evident that Sommerfeld's front velocity definition is inadequate, we assume that a signal $s_i(t)$ with the corresponding Fourier transform

$$\hat{s}_i(\omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s_i(t)\, e^{-j\omega t} \mathrm{d}t \tag{1}$$

is applied to the input at $x = 0$ of a transmission line with the transfer function $\hat{h}(\omega)$. We further assume without loss of generality that the input signal is normalized so that the total energy $\mathrm{E}\{s_i(t)\}$ is unity and consequently that equation

$$\mathrm{E}\{s_i(t)\} := \int_{-\infty}^{+\infty} s_i(t)^2\, \mathrm{d}t = 1 \tag{2}$$

is satisfied.

Let the transfer function of the transmission line be

$$\hat{h}(\omega) = e^{-j\omega x / v_p(\omega)}, \tag{3}$$

with $v_p(\omega)$ being the phase velocity at a given angular frequency ω and x being the location of the measurement point at the transmission line. As can be seen, this transfer function has unity gain for all frequencies and thus represents an all-pass filter. Note that for physically reasonable phase velocities, condition $v_p(\omega) = v_p(-\omega)$ must apply.

The effect of the transmission line on the input signal can be obtained by applying the inverse Fourier transform to the product of the signal spectrum $\hat{s}_i(\omega)$ and the transfer function $\hat{h}(\omega)$, i.e., by calculating

$$s_o(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{s}_i(\omega)\, \hat{h}(\omega)\, e^{j\omega t} \mathrm{d}\omega. \tag{4}$$

https://doi.org/10.32388/7QNDLJ

By substituting the transfer function (3) into (4), we obtain the output signal

$$s_o(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{s}_i(\omega)\, e^{j\,\omega\,(t - x/v_p(\omega))} \mathrm{d}\omega \tag{5}$$

of the transmission line for the applied input signal $s_i(t)$. As equation (5) clearly shows, the output signal $s_o(t)$ consists simply of running waves superimposed and weighted with the signal spectrum $\hat{s}_i(\omega)$ of the input signal $s_i(t)$. Thus, equation (5) is consistent with intuitive understanding.

III. Energy conservation

For the following it is important that the overall energy of the output signal $s_o(t)$ is preserved at any location $x > 0$. This aspect can be easily shown by using Parseval's theorem, which gives the equation

$$\mathrm{E}\{s_o(t)\} = \int_{-\infty}^{+\infty} |\hat{s}_i(\omega)\, \hat{h}(\omega)|^2 \,\mathrm{d}\omega = \int_{-\infty}^{+\infty} |\hat{s}_i(\omega)|^2\, |\hat{h}(\omega)|^2 \,\mathrm{d}\omega. \tag{6}$$

Because $|\hat{h}(\omega)| = 1$, we obtain

$$\mathrm{E}\{s_o(t)\} = \int_{-\infty}^{+\infty} |\hat{s}_i(\omega)|^2 \,\mathrm{d}\omega = \mathrm{E}\{s_i(t)\} = 1. \tag{7}$$

Consequently, a transmission line with the transfer function (3) does not change the energy of the signal.

IV. Specific phase velocity function

Now we can show that Sommerfeld's definition of frontal velocity is inadequate. For this purpose let us assume a specific and very simple phase velocity function:

$$v_p(\omega) = \begin{cases} u, & |\omega| \le \omega_u \\ c, & \text{otherwise.} \end{cases} \tag{8}$$

Herein, all phase velocities for $|\omega| \le \omega_u$ are equal to u but beyond that are equal to c.

Substituting the phase velocity function (8) into equation (5) yields the output signal

$$s_o(t) = \alpha\left(t - \frac{x}{u}\right) + \beta\left(t - \frac{x}{c}\right) - \alpha\left(t - \frac{x}{c}\right), \tag{9}$$

with the auxiliary functions α and β defined by

$$\alpha(\zeta) := \frac{1}{\sqrt{2\pi}} \int_{-\omega_u}^{+\omega_u} \hat{s}_i(\omega)\, e^{j\,\omega\,\zeta} \,\mathrm{d}\omega \tag{10}$$

and

$$\beta(\zeta) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{s}_i(\omega)\, e^{j\,\omega\,\zeta} \,\mathrm{d}\omega. \tag{11}$$

As is immediately apparent,

$$\beta(\zeta) = s_i(\zeta). \tag{12}$$

Furthermore, the output signal $s_o(t)$ can be resolved into a sum of two parts

$$s_o(t) = s_u(t) + s_c(t) \tag{13}$$

with

$$s_u(t) := \alpha\left(t - \frac{x}{u}\right) \tag{14}$$

and

$$s_c(t) := s_i\left(t - \frac{x}{c}\right) - \alpha\left(t - \frac{x}{c}\right). \tag{15}$$

As expected, the output signal $s_o(t)$ apparently consists of two components moving at velocities u and c, respectively. In the following, we examine the relevance of the components depending on the type of the input signal.

V. Dirac delta function as input

First, we study the properties of the transmission line defined in section IV when we apply the idealized impulse

$$s_i(t) = \sqrt{\delta(t)}. \tag{16}$$

As can be directly seen, constraint (2) is valid for this signal, i.e., the total energy of the signal is 1.

As shown in section IV, the output signal at any location x is given by equation (13), which indicates that the signal consists of two parts: (14) and (15). For input signal (16), we obtain the output signal

$$s_o(t) = \alpha\left(t - \frac{x}{u}\right) + \sqrt{\delta\left(t - \frac{x}{c}\right)} - \alpha\left(t - \frac{x}{c}\right). \tag{17}$$

Consequently, for the signal power

$$s_o(t)^2 = \delta\left(t - \frac{x}{c}\right) + 2\, s_r(t)\, \sqrt{\delta\left(t - \frac{x}{c}\right)} + s_r(t)^2 \tag{18}$$

holds with

$$s_r(t) := \alpha\left(t - \frac{x}{u}\right) - \alpha\left(t - \frac{x}{c}\right). \tag{19}$$

As shown in section III, the transmission line does not affect the energy of the signal. Therefore, the integral of the output signal power $s_o(t)^2$ is equal to 1. Moreover, the integral of $\delta\left(t - \frac{x}{c}\right)$ is 1 as well. Therefore, we get

$$\int_{-\infty}^{+\infty} 2\, s_r(t)\, \sqrt{\delta\left(t - \frac{x}{c}\right)} \mathrm{d}t + \int_{-\infty}^{+\infty} s_r(t)^2 \,\mathrm{d}t = 0. \tag{20}$$

The first integral in this equation can be rearranged, because

$$\int_{-\infty}^{+\infty} 2\, s_r(t)\, \sqrt{\delta\left(t - \frac{x}{c}\right)} \mathrm{d}t = \int_{-\infty}^{+\infty} \frac{2\, s_r(t)}{\sqrt{\delta\left(t - \frac{x}{c}\right)}}\, \delta\left(t - \frac{x}{c}\right) \mathrm{d}t. \tag{21}$$

However, the right-hand part, provided that $s_r(x/c) < \infty$, can be only 0, because as a result of the sifting property of the Dirac function, we obtain

$$\int_{-\infty}^{+\infty} \frac{2\, s_r(t)}{\sqrt{\delta\left(t - \frac{x}{c}\right)}}\, \delta\left(t - \frac{x}{c}\right) \mathrm{d}t = \frac{2\, s_r(\frac{x}{c})}{\sqrt{\delta\,(0)}} = 0. \tag{22}$$

But this means that the left term in equation (20) is zero and this in turn has the consequence that also

$$\int_{-\infty}^{+\infty} s_r(t)^2 \,\mathrm{d}t = 0 \tag{23}$$

must apply.

With equation (23) further follows that the signal $s_r(t)$ must be zero for any t, since the integral of an positive function can only be zero if the function is zero everywhere in the area of integration. If we now substitute $s_r(t) = 0$ into equation (18), we get

$$s_o(t)^2 = \delta\left(t - \frac{x}{c}\right), \tag{24}$$

i.e.

$$s_o(t) = s_i\left(t - \frac{x}{c}\right). \tag{25}$$

Consequently, the energy of the output signal is completely carried by the impulse, which propagates with velocity c. It is remarkable that this result even holds independently of the value of the parameter ω_u in the definition of the phase velocity (8).

This finding appears to confirm the conclusions reached by A. Sommerfeld [5]. The effect can also be explained and interpreted by examining the spectrum of the Dirac impulse: because its spectrum is constant and independent of frequency, the energy of the signal in any low-frequency range is always infinitesimally small relative to the overall energy.

Nevertheless, arguing that signal fronts therefore cannot propagate faster than the impulse response would be a mistake. The next section makes this evident by studying an input signal with nonzero temporal duration.

VI. Ideal rectangular pulse as input

Now we analyze an input signal with a nonzero temporal duration, i.e.,

$$s_i(t) = \frac{1}{\sqrt{\tau}}\left(\Theta(t) - \Theta(t-\tau)\right). \tag{26}$$

Here, Θ is the Heaviside step function. As can be verified, the condition (2) holds for this signal, i.e., the overall energy is 1. Of note, for $t < 0$ the signal is exactly 0, and for $t = 0$, it has an infinitely steep slope. The signal is therefore completely causal and contains infinitely high frequencies.

The Fourier transform $\hat{s}_i(\omega)$ can be easily calculated:

$$\begin{aligned}\hat{s}_i(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} s_i(t)\, e^{-j\omega t}\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi\tau}}\int_{0}^{\tau} e^{-j\omega t}\mathrm{d}t \\ &= \frac{j\left(e^{-j\omega\tau} - 1\right)}{\sqrt{2\pi\tau}\,\omega}.\end{aligned} \tag{27}$$

The term $\alpha(\zeta)$, defined by equation (10), can we obtain by substituting equation (27) and computing the integral. We get

$$\alpha(\zeta) = \frac{1}{\pi\sqrt{\tau}}\left(\mathrm{Si}(\omega_u\,\zeta) + \mathrm{Si}(\omega_u(\tau - \zeta))\right). \tag{28}$$

The output signal $s_o(t)$ is now determined for all locations x, because all terms in equation (13) are now known and we have

$$s_o(t) = \alpha\left(t - \frac{x}{u}\right) + s_i\left(t - \frac{x}{c}\right) - \alpha\left(t - \frac{x}{c}\right). \tag{29}$$

Figure 1 shows the waveform of the signal power $s_o(t)^2$ for example parameters at different distances x from the input of the transmission line.

As can be clearly seen from equation (13) and in figure 1, the energy of the signal part, which moves with velocity u, does not disappear in this case. For the overall energy of the output signal, the following applies:

$$\mathrm{E}\{s_o(t)\} = \mathrm{E}\{s_u(t)\} + \mathrm{E}\{s_c(t)\} + 2\int_{-\infty}^{+\infty} s_u(t)\, s_c(t)\,\mathrm{d}t. \tag{30}$$

The first term $\mathrm{E}\{s_u(t)\}$ represents the energy of the signal part that propagates entirely at velocity u. Because of Parseval's theorem, the equation

$$\mathrm{E}\{s_u(t)\} = \int_{-\infty}^{+\infty} |\hat{s}_u(\omega)|^2\,\mathrm{d}\omega \tag{31}$$

holds true. As can be seen from equations (10) and (14),

$$\hat{s}_u(\omega) = \left(\Theta(\omega + \omega_u) - \Theta(\omega - \omega_u)\right)\,\hat{s}_i(\omega)\, e^{-j\,\omega\, x/u}, \tag{32}$$

i.e.,

$$\mathrm{E}\{s_u(t)\} = \int_{-\omega_u}^{+\omega_u} |\hat{s}_i(\omega)|^2\,\mathrm{d}\omega. \tag{33}$$

From this, it follows, by substituting equation (27)

$$\begin{aligned}\mathrm{E}\{s_u(t)\} &= \int_{-\omega_u}^{+\omega_u} \frac{\left(e^{-j\,\omega\tau} - 1\right)}{\sqrt{2\pi\,\tau}\,\omega}\,\frac{\left(e^{j\omega\tau} - 1\right)}{\sqrt{2\pi\,\tau}\,\omega}\,\mathrm{d}\omega \\ &= \int_{-\omega_u}^{+\omega_u} \frac{1 - \cos(\omega\,\tau)}{\pi\,\omega^2\,\tau}\,\mathrm{d}\omega.\end{aligned} \tag{34}$$

Evaluating the integral yield

$$\mathrm{E}\{s_u(t)\} = \frac{2\,\left(\cos(\tau\,\omega_u) + \tau\,\omega_u\,\mathrm{Si}(\tau\,\omega_u) - 1\right)}{\pi\,\tau\,\omega_u}. \tag{35}$$

As equation (35) reveals, the energy of this signal part depends only on the product $\tau \cdot \omega_u$. For $\tau \to \infty$, $\mathrm{E}\{s_u(t)\}$ becomes equal to 1. Therefore, in this case, all the energy of the signal moves with velocity u, and the part with velocity c disappears. Of note, by increasing τ, as can be seen from equation (26), the pulse only becomes wider. The infinitely steep slope at the input of the transmission line at time $t = 0$ remains unchanged. Furthermore, it remains valid that the signal is exactly 0 for $t < 0$.

In addition, for finite τ, i.e., for true rectangular signals, a major part of the energy can be contained in the signal part that propagates with velocity u. For example, for $\tau\,\omega_u = 1000$ the energy $\mathrm{E}\{s_u(t)\}$ is approximately 0.99937. Even for $\tau\,\omega_u = 2\,\pi$, there is still a significant amount of energy in the signal component $s_u(t)$, because in this case $\mathrm{E}\{s_u(t)\} \approx 0.90282$.

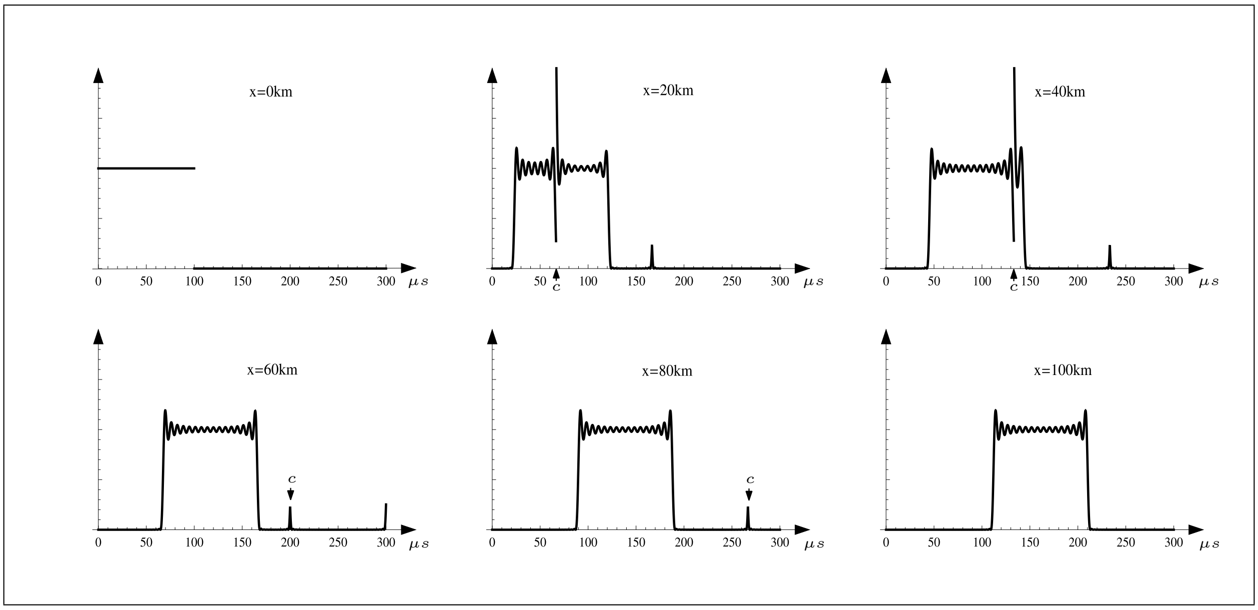

Fig. 1. The figure shows the waveform of the energy $s_o(t)^2$ of the signal at different distances x from the input of the transmission line. In this example, $u = 3\,c$, $\tau = 100\,\mu$s and $\omega_u = 1\,$MHz. As can be seen, the rectangular pulse propagates with velocity $3\,c$. The high-frequency components, in contrast, have only the velocity c and form small glitches, which are increasingly left behind with increasing distance (as an aid, the locations that can be reached by a signal with velocity c are marked by small arrows). The low frequency part of the signal $s_u(t)$ clearly does not have an infinitely steep slope. However, this does not mean that there is no clearly recognizable front.

For $\tau \to 0$, however, $\mathrm{E}\{s_u(t)\}$ becomes 0, and we again have the case studied in section V.

VII. Conclusions

As shown in the previous two sections, the transmission line in section IV behaves strangely. If we were to attempt to measure the impulse response, we would find that the impulse propagates with velocity c along the transmission line. We would not notice that there is an additional part that propagates with velocity u, because the energy of that part is close to 0.

However, if we were to measure the step response, we would observe that although the front loses steepness because of the low-pass filtering, the front needs only the time x/u and not the time x/c to reach the location x. However, we would not notice that in addition, there is a part that propagates with velocity c, because that part contains practically no energy.

Thus, one might be inclined to state that the impulse response propagates with velocity c, but the step response propagates with velocity u. This conclusion seems to diretly contradict the rule that the step response can be represented as the convolution of the unit step or Heaviside function with the impulse response. Therefore, several questions arise:

1) Is this effect an artifact of the non-causality of the studied transmission line?
2) How can this paradox be resolved?
3) How should we define the front velocity correctly?

Question 1 can be answered quickly by imagining the total transmission line as a parallel network of two separate transmission lines, as shown in figure 2. In this model, one transmission line transmits the low-frequency part of the input signal, and the other transmits the high-frequency part by placing *causal* low-pass and high-pass filters before each line. Subsequently, the delay behavior is modeled with ideal delay elements with delay constants x/u and x/c. These ideal delay elements are also causal, because they do not produce any dispersion.

As can easily be seen, the complete transmission line is causal. Nevertheless, nothing changes the fact that the limit of the phase velocity for $\omega \to \infty$ is c and that the impulse response propagates with velocity c, while the step response moves mainly with velocity u. For causal filters, the existence of a mixed region where both transmission lines are conductive is not important for this argument.

To further illustrate this statement, we consider the two most simple known causal filters, namely resistor-capacitor circuits. The transfer function of the low-pass filter is with $\tau := RC$

$$\hat{h}_L(\omega) = \frac{1}{1 + j\,\omega\,\tau}. \tag{36}$$

The transfer function of the complementary high-pass filter is

$$\hat{h}_H(\omega) = \frac{j\,\omega\,\tau}{1 + j\,\omega\,\tau}. \tag{37}$$

The total transfer function corresponding to figure 2 is then

$$\hat{h}(\omega) = \hat{h}_L(\omega)\, e^{-j\,\omega\,\frac{x}{u}} + \hat{h}_H(\omega)\, e^{-j\,\omega\,\frac{x}{c}}. \tag{38}$$

It is not difficult to calculate the corresponding step response, that means the output signal

$$s_o(t) = \Theta\left(t - \frac{x}{u}\right)\left(1 - e^{-\frac{1}{\tau}\left(t - \frac{x}{u}\right)}\right) + \Theta\left(t - \frac{x}{c}\right) e^{-\frac{1}{\tau}\left(t - \frac{x}{c}\right)} \tag{39}$$

of the input signal

$$s_i(t) = \Theta(t). \tag{40}$$

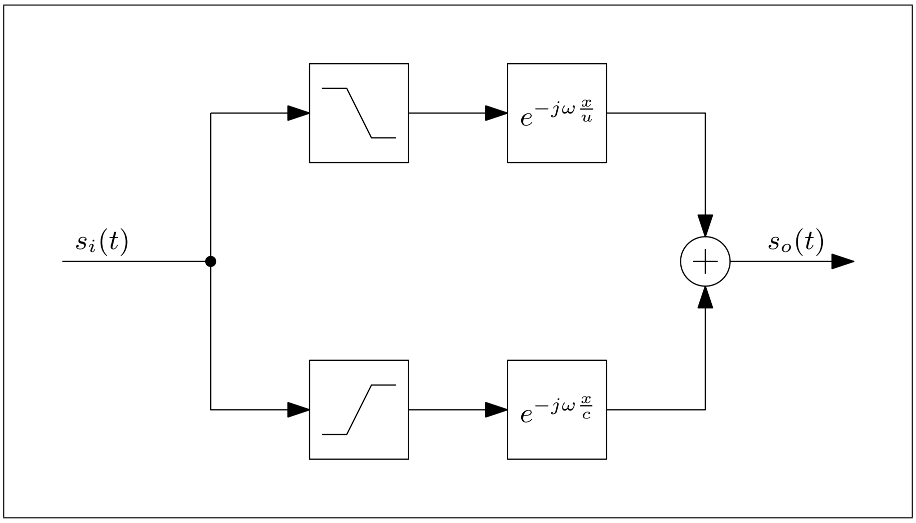

Fig. 2. This transmission line is causal, because all components are causal: the upper channel is a series of a causal low-pass filter and a causal transmission line in which all phase velocities have the same value, u. The lower channel is a series of a causal high-pass filter and a causal transmission line in which all phase velocities have the value c.

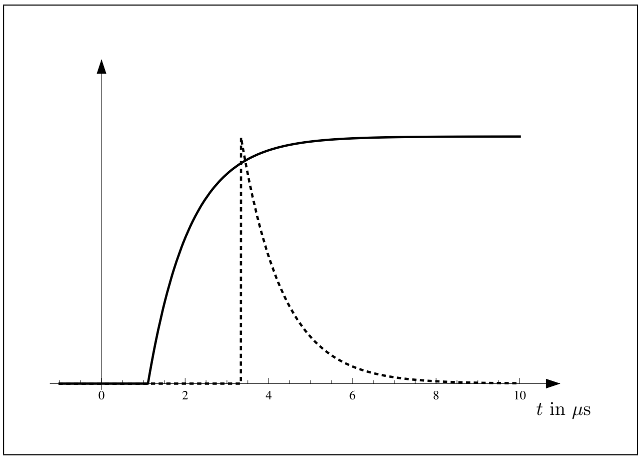

Fig. 3. Parts of signal $s_o(t)$ in equation (39) at x = 1000m for $\tau = 1\mu$s and $u = 3\,c$. It is obvious that the signal is causal, but the step response moves faster than the impulse response.

As can be easily seen, $s_o(t)$ is perfectly causal and consists of two parts with two fronts, namely

- one with an infinitely steep slope which moves with velocity c (Fig. 3, dotted line) and
- one which rises only with $1/\tau$ but propagates with velocity u (Fig. 3, solid line).

But the limit of the phase velocity for $\omega \to \infty$ is obviously equal to c. Nevertheless, there is a signal part that is propagating at velocity u independently of this limit. This, however, demonstrates once more that Sommerfeld's front velocity definition is definitely inadequate and incorrect, since the term "front" suggests that this part is in *any case* the fastest. For $u > c$, however, this is not true.

Figure 2 also answers question 2: the apparent paradox can be resolved because in measuring the impulse response, essentially only the properties of the lower channel are effective, whereas in measuring the step response, essentially only the upper channel has an effect. Mathematically, however, the impulse response is given by equation (17), which also contains components that move with velocity u, even if they are not practically measurable. In conclusion, measuring only the impulse response or only the step response may be insufficient, because important information may be missing in both.

Finally question 3 remains regarding how to define the front velocity. First of all, it has become obvious that the phase velocity for infinitely high frequency represents *not* the upper limit for the propagation speed of signal fronts. Instead, it is only an impulse response velocity and does not necessarily provide information about the speed at which the step response or information propagates. But since the front speed is supposed to be the speed at which the fastest signal components propagate, the conclusion can only be that the front velocity is equal to the fastest phase velocity, provided that the phase velocity function is continuous in the vicinity of the corresponding frequency.

VIII. Acknowledgments

The author thanks Max Minh Tran for his support and comments.

References

[1] J. L. Leander, "On the relation between the wavefront speed and the group velocity concept," *J. Acoust. Soc. Am*, vol. 100, no. 6, 1996. [Online]. Available: https://doi.org/10.1121/1.417249

[2] P. W. Milonni, *Fast light, slow light and left-handed light*. CRC Press, 2004.

[3] R. Chiao, "Superluminal phase and group velocities: A tutorial on sommerfeld's phase, group, and front velocities for wave motion in a medium, with applications to the "instantaneous superluminality" of electrons," 2011. [Online]. Available: https://arxiv.org/abs/1111.2402

[4] W. Withayachumnankul, B. M. Fischer, B. Ferguson, B. R. Davis, and D. Abbott, "A systemized view of superluminal wave propagation," *Proceedings of the IEEE*, vol. 98, no. 10, pp. 1775–1786, 2010.

[5] A. Sommerfeld, "Über die Fortpflanzung des Lichtes in dispergierenden Medien," *Annalen der Physik*, vol. 44, no. 4, 1914.

[6] T. Sauter, "Superluminal signals: an engineer's perspective," *Physics Letters A*, 2001. [Online]. Available: https://doi.org/10.1016/S0375-9601(01)00183-9

[7] G. B. Malykin and E. A. Romanets, "Superluminal motion (review)," *Optics and Spectroscopy*, vol. 112, no. 6, pp. 920–934, 2012.

[8] S. Hrabar, I. Krois, I. Bonic, and A. Kiricenko, "Ultra-broadband simultaneous superluminal phase and group velocities in non-foster epsilon-near-zero metamaterial," *Appl. Phys. Lett.*, vol. 102, no. 054108, 2013.

[9] M. S. Kulya, V. A. Semenova, V. G. Bespalov, and N. V. Petrov, "On terahertz pulsed broadband gauss-bessel beam free-space propagation," *Scientific Reports*, vol. 8(1):1390, 2018.